THE UNIVERSITY *of* EDINBURGH

UNIVERSITY OF EDINBURGH

DOCTORAL THESIS

Enhanced Raman spectroscopy as a tool for the improvement of cervical cancer screening: a comparison of global and targeted approaches

---

Author:
Lana WOOLFORD

Supervisors:
Prof. C. Simon HERRINGTON
Prof. Duncan GRAHAM
Prof. Ted HUPP

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Centre for Doctoral Training in Optical Medical Imaging
Department of Chemistry

July 16, 2020

This page originally contained the Declaration of Authorship and is intentionally left blank as per University guidelines.

Abstract

Enhanced Raman spectroscopy as a tool for the improvement of cervical cancer screening: a comparison of global and targeted approaches

by Lana WOOLFORD

Half a million women worldwide are diagnosed with cervical cancer annually, with 87% of mortalities occurring in developing economies. HPV vaccination has reduced disease incidence significantly but must be supplemented with regular cervical screening. In established practice, cells from the cervix are fixed, stained and observed manually using an optical microscope. Inter-observer subjectivity and false diagnoses are a relatively frequent occurrence.

The use of molecular pathology allows objective diagnoses to be made based on changes to the molecular content of cells and tissue with disease. This quantitative, objective and automated pathological analysis could be integrated with existing services, as well as creating potential for economically deprived regions. Molecular pathology through spectroscopic methods has shown great promise. For Raman spectroscopy, this is broadly categorised into global comparison or targeted biomarker approaches.

In the global paradigm, a direct comparison of wavelength modulated with standard Raman spectroscopy showed marked improvement of sensitivity and specificity in fixed cell classification. Intracellular sampling location studies indicated that the technique is more robust for fixed than live cells. A functional nanoparticle bioconjugate for the established marker p16 was developed for targeted studies following optimisation of covalent, passive and bioconjugation methods. This was applied to fixed smear analogues, both lysed and intact, in lateral flow assay and SERS mapping contexts respectively. Considerations of hybrid SERS-SRS and global-targeted imaging were also explored, combining key vibrational modes and labelling moieties from both arms of the study.

After consideration of the clinical context, the SERS-active paper-based dipstick assay was selected for further development. Novel binding peptides for p16 capture were discovered and validated, alternative consensus motifs found through next-generation sequencing and SERS-based quantification demonstrated to be sufficiently sensitive to detect relevant cellular analogues. The investigations described provide both a roadmap for further development of multiple Raman diagnostic tools in a clinical context and the foundational components of a point-of-care, inexpensive test for p16 which could streamline and broaden access to cervical screening.

Acknowledgements

I would like to thank, first of all, my supervisors - Professors Simon Herrington, Duncan Graham and Ted Hupp - for their support, advice and good humour throughout the project. I think we achieved something genuinely interdisciplinary, which is no mean logistical feat between multiple universities, campuses, labs and ways of thinking! All of that was made possible with funding from the Medical Research Council, Engineering and Physical Sciences Research Council, and Cancer Research UK via the Centre for Doctoral Training in Optimal Medical Imaging, OPTIMA CDT.

OPTIMA has played a very significant role in my development as a scientist and as a person. Beyond studying alongside fellow CDT students and giving me a more rounded outlook, it has provided me with an enduring group of friends who seem to share not just a love of science but a passion for using our skills to make the world a better place. Thanks to the unsung heroes of the programme for making everything possible - Dr Jean O'Donoghue, Samantha Brown, Gillian Neeson and Brittany Bovenzy. Special thanks go to Dr Kirsty Ross for introducing me to the joys of public engagement with science, and going the extra mile in supporting my personal and professional development. Your encouragement landed me a fantastic placement with the talented Claudia Stocker at Vivid Biology Ltd, won me an award and helped me find my career.

Enormous thanks to my host lab family, the Hupp/Ball groups, for all of your technical help and friendship. I'd not so much as made a buffer when I stepped into the Edinburgh Cancer Research Centre in 2015. I would like to give special thanks to my bench buddy and general wet-lab Jedi master Maria Gomez-Herranz, for your enduring patience and philosophical chats while I learned the ropes of proteomics. Thanks also to 'ECRC mum' Jan Irvine for looking after us all, and to fellow students Andrea Coates and David Hay for keeping me sane!

Finally, I would like to thank my mum, dad, sister and wider family for your ongoing love and support, especially during low points. I couldn't have done this without your help and understanding. And then of course there are my adopted family members from Old Saint Paul's Church - you welcomed me, kept me sane and introduced me to this amazing country.

I cannot thank you all enough.

Contents

Declaration of Authorship **iii**

Abstract **v**

Acknowledgements **vii**

1 Introduction **1**
1.1 Cervical cancer . . . 2
1.1.1 Burden of disease . . . 2
1.1.2 Mechanisms of disease . . . 4
1.1.3 Clinical presentation of cervical cancer . . . 6
1.1.4 Current clinical practice . . . 6
1.1.5 p16 as a marker of cervical cancer and pre-cancer . . . 9
1.1.6 Alternative biomarkers for triage of hrHPV primary screening . . . 10
1.2 Molecular pathology . . . 14
1.2.1 Global and targeted approaches . . . 14
1.2.2 Advantages and disadvantages of molecular pathology . . . 15
1.2.3 Current examples in clinical practice . . . 16
1.3 Raman spectroscopy as a tool for molecular pathology . . . 17
1.3.1 Fundamentals of Raman spectroscopy . . . 17
1.3.2 Raman spectroscopy in molecular pathology . . . 20
1.3.3 Raman spectroscopy in cervical cancer diagnosis . . . 22
1.4 Exploring Raman spectroscopy as a tool for improved cervical screening practice . . . 26

2 General materials and methods **27**
2.1 Cell culture . . . 27
2.1.1 Primary cells . . . 27
2.1.2 Established cell lines . . . 28
2.2 Production of his-tagged p16 . . . 29
2.2.1 Transfection . . . 29

2.2.2 Purification . . . 30
2.3 Western Blotting . . . 34
2.3.1 Cell lysing and protein quantification . . . 34
2.3.2 Gel electrophoresis and imaging . . . 36
2.4 Immunofluorescence . . . 37
2.4.1 Cell and slide preparation . . . 37
2.4.2 Fluorescence imaging and analysis . . . 37
2.5 Enzyme-linked immunosorbent assay (ELISA) . . . 38
2.6 Paper-based immunoassays . . . 39
2.6.1 In-house dipstick assays . . . 39
2.6.2 Commercial conjugation assays . . . 40
2.7 Nanoparticle blotting . . . 41
2.8 SERS mapping . . . 42

3 Wavelength modulated Raman spectroscopy for cell classification 43
3.1 Introduction . . . 43
3.2 Concept: wavelength modulation . . . 44
3.3 Materials and methods . . . 45
3.3.1 Cell sample preparation . . . 45
3.3.2 Spectral acquisition . . . 45
3.3.3 Spectral processing and data analysis . . . 46
3.3.4 Cross-validation of spectral classification . . . 48
3.4 Comparison of human papillomavirus (HPV) status . . . 49
3.5 Comparison of transformation status . . . 53
3.5.1 Comparison of standard and modulated spectral classification by PCA . . . 53
3.5.2 Pairwise analysis of cell lines . . . 53
3.6 Considerations of sampling location . . . 61
3.7 Spectral variation due to sample storage . . . 67
3.8 Analysis of significant vibrational modes . . . 69
3.9 Conclusions . . . 71

4 Development of antibody-conjugated nanoparticles 73
4.1 Introduction . . . 73
4.2 Concept: SERS and nanoparticle conjugates . . . 79
4.3 Materials and methods . . . 81
4.3.1 Nanoparticle production . . . 81
4.3.2 Raman spectroscopy . . . 81

4.3.3 Covalent conjugation . . . 82
4.3.4 Passive conjugation . . . 82
4.3.5 Bioconjugation . . . 82
4.4 Production of Raman-active nanoparticles . . . 83
4.4.1 Gold nanoparticle characterisation . . . 83
4.4.2 Investigation of Raman reporters . . . 84
4.4.3 Optimisation of BPE-AuNP stability and signal . . . 88
4.4.4 Effect of buffers on signal and stability . . . 90
4.5 Antibody characterisation . . . 92
4.5.1 Stability and sensitivity of commercial antibodies . . . 92
4.5.2 Antibody limit of detection . . . 93
4.6 Covalent conjugation of antibody . . . 95
4.6.1 Covalent conjugation scheme . . . 95
4.6.2 Comparison of EDC/NHS methods . . . 95
4.6.3 Optimisation of reagent concentration . . . 102
4.6.4 Variability of binding functionality . . . 107
4.7 Passive conjugation of antibody . . . 111
4.7.1 Passive conjugation scheme . . . 111
4.7.2 Optimisation of antibody coverage and surface competition . . . 112
4.7.3 Effect of pH on passive conjugates . . . 113
4.7.4 Passive conjugates with reduced reporter coverage . . . 115
4.7.5 Consideration of passive conjugate blocking on protein interaction 117
4.8 Bioconjugation of antibody . . . 121
4.8.1 Bioconjugation scheme . . . 121
4.8.2 Optimisation of Protein G′ coverage . . . 121
4.8.3 Conjugation of antibody on Protein G′ coated particles . . . 125
4.8.4 Storage and signal considerations . . . 126
4.9 Conclusions . . . 132

5 Quantification of p16 in whole cells 135
5.1 Introduction . . . 135
5.2 Concept: ratiometric imaging approaches . . . 136
5.3 Materials and methods . . . 137
5.3.1 Cell incubation with conjugates . . . 137
5.3.2 Fluorescent antibody labelling . . . 138
5.4 Production of a physicochemical control particle . . . 139
5.4.1 Binding validation . . . 139

5.4.2 Physical characteristics . . . 140
5.5 Ratiometric conjugates in whole cells . . . 144
5.5.1 Preliminary optimisation of conjugate incubation . . . 144
5.5.2 Development of conjugate incubation and ratiometric method . . 146
5.5.3 Establishing a signal processing technique . . . 149
5.6 Optimising fluorescent labelling as an analogue for alkyne-tagged antibodies . . . 156
5.6.1 Optimisation of NHS coupling using commercial labels . . . 157
5.6.2 Lifetime and storage considerations . . . 159
5.6.3 Comparison of commercial and in-house conjugation using immunofluorescence . . . 159
5.7 Conclusions . . . 165

6 Development of a sensitive paper-based SERS diagnostic device 167
6.1 Introduction . . . 167
6.2 Concept: paper-based diagnostic devices . . . 172
6.3 Materials and methods . . . 175
6.3.1 Phage-displayed peptide biopanning . . . 176
6.3.2 ELISAs . . . 180
6.3.3 Phage samples for next-generation sequencing (NGS) . . . 184
6.3.4 NGS data analysis workflow . . . 187
6.4 Peptide discovery for p16 protein and antibody capture . . . 188
6.4.1 Optimisation of the biopanning process . . . 188
6.4.2 Discovery of peptide sequences . . . 190
6.4.3 Affinity of synthetic peptides by ELISA . . . 194
6.4.4 Amino acid analysis of peptide sequences . . . 199
6.5 Peptide activity with cell models . . . 201
6.5.1 Quantification of p16 in cell lines . . . 201
6.5.2 Peptide binding efficacy with cell lysate . . . 202
6.6 Next generation sequencing of phage populations . . . 204
6.6.1 Large-scale sequencing of phage populations . . . 204
6.6.2 Analysis of large-scale peptide sequencing . . . 205
6.7 SERS activity of paper-based devices . . . 208
6.7.1 Choosing a substrate architecture . . . 208
6.7.2 Raman response of assembled dipsticks . . . 211
6.7.3 Detection limits of commercial dipsticks . . . 212
6.8 Conclusions . . . 216

7 Conclusion and suggestions for further work **219**

8 Appendix A: Published research **225**

9 Appendix B: SERS-SRS imaging **237**
9.1 Materials and methods 237
9.1.1 Stimulated SERS mapping 237
9.2 Consideration of stimulated SERS imaging 238

Bibliography **241**

List of Figures

1.2 HPV genome and cell cycle 4
1.3 HPV oncogenic pathway 6
1.4 Expression of HPV genes during viral replication 7
1.5 Cervical lesion progression 12
1.6 LSIL and HSIL cytology and histopathology 13
1.7 Former UK clinical pathway for cervical screening 13
1.8 Cervical cancer patient management through molecular pathology . . . 17
1.9 Spontaneous Raman spectroscopy and biological spectra 18

2.1 GFP transfection optimisation 32
2.2 Nickel plate capture of His-p16 33
2.3 Western blot with different lysis buffers 35
2.4 Tiled immunofluorescence 38
2.5 Initial p16 ELISA 39
2.6 Constructed dipstick assays 40
2.7 Dipstick assay method 41
2.8 Nanoparticle blotting on wax-printed assays 42

3.1 WMRS cavity slide assembly 46
3.2 WMRS optical setup 47
3.3 LOOCV method 49
3.4 HPV standard PCA plot 51
3.5 HPV modulated PCA plot 52
3.6 HPV average Raman spectra 52
3.7 Transformation status average spectra 54
3.8 Transformation status PCA plots 55
3.9 Pairwise sensitivity values 57
3.10 Pairwise specificity values 58
3.11 Significant standard spectral regions 59
3.12 Significant modulated spectral regions 60

3.13 Adhered and suspended cell morphology 61
3.14 SiHa PCA plot without sampling location 62
3.15 SiHa PCA plot expanded by intracellular location 63
3.16 CaSki PCA plot expanded by intracellular location 64
3.17 HaCaT PCA plot expanded by intracellular location 65
3.18 NHEK PCA plot expanded by intracellular location 66
3.19 Raman spectral variation with storage time 68
3.20 PCA plot of stored cell samples 68

4.1 Components of a typical SERS nanotag 74
4.2 Examples of colloidal nanostructures 75
4.3 Extinction spectra of different nanoparticle types 76
4.4 Antibody orientation 79
4.5 Localised surface plasmon resonance 79
4.6 Method for covalent conjugation 83
4.7 Gold nanoparticle size distribution 84
4.8 Raman reporters 86
4.9 Reporter absorbance spectra 87
4.10 Optimising BPE concentration 88
4.11 Raman signal of BPE 89
4.12 Limit of Raman detection of BPE 90
4.13 BPE Raman spectra in different buffers 91
4.14 BPE Raman spectra after swapping buffers 92
4.15 Choice of commercial antibody to p16 93
4.16 Sensitivity ELISAs for BD Pharmingen mAb to p16 94
4.17 Covalent conjugation scheme 96
4.18 Covalent conjugation using EDC and sulfo-NHS 97
4.19 IgG1 antibodies with on-NP conjugation 99
4.20 Effect of antibody variation on absorbance (on-NP) 100
4.21 Effect of antibody variation on physical characteristics (on-NP) 100
4.22 Comparison of on- and off-NP conjugation 101
4.23 Absorbance characteristics of different reagent combinations 104
4.24 Absorbance comparison of different reagent variations 105
4.25 Commercial dipstick assay results for antibody-conjugated samples . . . 106
4.26 Dipstick assay results for covalent conjugates 106
4.27 Stability of conjugation methods and reagent variation 108
4.28 Dipstick assay binding of EDC/NHS conjugates 109

4.29 Raman spectra of conjugation method and reagent variation 110
4.30 Passive conjugation scheme 111
4.31 Conjugate stability with variation of antibody volume 114
4.32 Conjugate characteristic variation with antibody concentration 115
4.33 Variation in reporter-NP characteristics in triplicate 116
4.34 Extent of antibody coverage with variation of functional molecules . . . 117
4.35 Passive conjugation binding qualities 117
4.36 Conjugate blot-binding with variation in blocking agent 118
4.37 Interaction of antibody with BSA 119
4.38 Antibody interaction specificity and milk blocked membranes 120
4.39 Alternative blocking agents and dipstick assay buffer 120
4.40 Passive and bioconjugate blotting signals 121
4.41 Optimisation of Protein G' coverage 123
4.42 Variation of physical characteristics with antibody coverage 124
4.43 Conjugate blotting signal with antibody concentration 126
4.44 Raman spectral variation with conjugation and storage buffer 128
4.45 Characteristics of bioconjugates with higher reporter concentration . . . 130
4.46 Raman spectra of higher concentration BPE-AuNPs 131
4.47 Optimised targeted bioconjugate in triplicate 131
4.48 Recommended workflow for conjugate optimisation 134

5.1 Binding validation of target and dummy conjugates 141
5.2 PPY titration characteristics 141
5.3 PPY samples in triplicate 142
5.4 Reference ladder of bulk ratiometric samples 143
5.5 SERS hotspot maps for fixed w/t SiHa and HaCaT 145
5.6 Reference spectra for target and dummy conjugates 147
5.7 SERS hotspot maps for w/t and transfected PreservCyt-fixed SiHa . . . 148
5.8 Raw hotspot spectral data for w/t and transfected fixed SiHa 149
5.9 Ratiometric SERS maps for w/t and transfected fixed SiHa 150
5.10 Ratiometric spectral data for w/t and transfected fixed SiHa 152
5.11 Reference spectra for ratiometric conjugates 153
5.12 Ratiometric reference data for spectral analysis 154
5.13 False-colour ratiometric maps 155
5.14 BADY-succinimide 156
5.15 Fluorophore conjugation with 10ml column 157
5.16 Absorbance spectra of conjugation optimisation samples 158

5.17 Immunofluorescence lifetime and storage . . . 160
5.18 Absorbance and structure of Alexa Fluor 488 . . . 161
5.19 Fluorescent conjugate absorbance profiles . . . 162
5.20 FITC intensity of SiHa w/t with in-house AlexaFluor 488 labelling . . . 162
5.21 Alexa Fluor 488-conjugated mAbs vs secondary-only control . . . 163
5.22 Fluorescence histograms for Alexa Fluor 488-conjugated mAbs . . . 164

6.1 Lateral flow publications . . . 168
6.2 Paper assay architecture . . . 170
6.3 p16 epitopes and ligands in paper assay format . . . 173
6.4 Schematic of the biopanning process . . . 174
6.5 Sequence map of M13KE bacteriophage . . . 177
6.6 Example of blue/white plaque titration . . . 180
6.7 Phage library insert sequence . . . 181
6.8 Amino acid groupings . . . 182
6.9 Primer design for Illumina sequencing . . . 185
6.10 Next generation sequencing . . . 187
6.11 NGS FastQC data . . . 188
6.12 Pooled and individual phage characterisation . . . 189
6.13 ELISA affinity of native AN phages . . . 193
6.14 ELISA affinity of native AB phages . . . 194
6.15 Peptide sequence summary . . . 195
6.16 ELISA affinity of synthetic peptides . . . 196
6.17 Control peptides . . . 198
6.18 Dilution curve of p16 against peptides . . . 199
6.19 Peptide sequence alignment . . . 200
6.20 p16 reference Western Blots . . . 202
6.21 Binding affinity of peptide to p16 in cell lysate . . . 203
6.22 Band visualisation of ready-to-sequence library . . . 205
6.23 Peptide motif from NGS analysis . . . 208
6.24 Paper-based SERS spectra . . . 209
6.25 Conjugate response in wax-printed assay format . . . 210
6.26 Conjugate response in assembled dipsticks . . . 211
6.27 Raman maps of commercial dipsticks . . . 213
6.28 SERS spectra of commercial dipsticks . . . 214
6.29 Limit of detection of commercial assays . . . 215

7.1 Strengths and suggested directions for Raman spectroscopic studies . . 224

9.1 SRS optical setup 238
9.2 SRS spectral sweep of SiHa 240

List of Tables

2.1 Cell lines cultured and supporting reagents . . . 28
2.2 Optimised transfection conditions . . . 30
2.3 Western blotting buffers . . . 34
2.4 Handcast SDS gel reagents . . . 35

3.1 Classification of true and false positive and negative samples . . . 49
3.2 Standard transformation status confusion matrix . . . 53
3.3 Modulated transformation status confusion matrix . . . 56
3.4 Classification accuracy of intracellular acquisition samples . . . 62
3.5 Intracellular sampling classification accuracy . . . 63
3.6 Intracellular sampling with reduced spatial sensitivity . . . 63
3.7 Stored samples confusion matrix . . . 68
3.8 Statistically significant Raman spectral regions for classification . . . 70

4.1 Common Raman reporters . . . 77
4.2 Nanoparticle characteristics . . . 85
4.3 Reporter-NP characteristics . . . 86
4.4 Reagent concentrations for covalent conjugation conditions . . . 103
4.5 Bioconjugate characteristics for initial antibody addition . . . 125
4.6 Conjugation buffer considerations for bioconjugates . . . 127
4.7 Optimisation of bioconjugate labelling . . . 129
4.8 Size of final bioconjugates . . . 129

5.1 Reagent parameters for ratiometric conjugates . . . 140
5.2 Absorbance characteristics of target and dummy conjugates . . . 143
5.3 Ratiometric conjugate labelling accuracy in fixed cells . . . 146
5.4 DOL for preliminary fluorophore conjugation . . . 158
5.5 DOL for optimised fluorophore conjugation . . . 161

6.1 Buffers and reagents for peptide discovery . . . 175
6.2 Buffers and reagents for phage PCR . . . 186

6.3 Initial AN peptide sequences 190
6.4 Initial AB peptide sequences 191
6.5 Antigen peptides by binding affinity 197
6.7 p16 expression in cell line models 202
6.8 Phage peptide insert sequences and frequencies 205

List of Abbreviations

%T	Total monomer concentration
AB	Antibody-binding peptide
AgNP	Silver nanoparticle
AN	Antigen-binding peptide
APS	Ammonium persulphate
ATP	p-aminothiophenol
AuNP	Gold nanoparticle
AZPY	4,4-azopyridine
BADY	N-4-(4-phenyl-1,3-butadiyn-1-yl)benzyl
BPE	1,2-di(4-pyridyl)ethylene
BPT	1,1′-biphenyl-4-thiol
BSA	Bovine Serum Albumin (Fraction V)
BSA	Bovine serum albumin (Fraction V)
BT	Benzothiol
C-C	Carbon-carbon single bond
CARS	Coherent anti-Stokes Raman spectroscopy
CCD	Charge-coupled device
CDK	Cyclin dependant kinase
CF	Correction factor
CH	Carbon-hydrogen single bond
CH_2	Methylene
CH_3	Methyl
CHO	Chinese Hamster Ovary
CMOS	Complementary metal-oxide semiconductor
C-O	Carbon monoxide bond
CO2	Carbon dioxide
COD	Coefficient of determination
CT-PEG	Carboxyl-thiol terminated PEG
DAPI	4′,6-diamidino-2-phenylindole
DCLS	Direct classical least squares

dH_2O	Ultrapure water
DLS	Dynamic light scattering
DMEM	Dulbecco's Modified Eagle Media
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOL	Degree of labelling
DTT	Dithiothreitol
DTTC	3,3′-Diethylthiatricarbocyanine
ECL	Enhanced chemiluminescence
EDC	N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride
EDTA	Ethylenediamine tetraacetic acid
EF	Enhancement factor
ELISA	Enzyme-linked immunosorbent assay
EtOH	Ethanol
FBS	Foetal bovine serum
Fc region	fragment crystallizable region
FITC	Fluorescein isothiocyanate
FWHM	Full width half maximum
GFP	Green fluorescent protein
HCl	Hydrochloric acid
HDI	Human Development Index
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HNO_3	Nitric acid
HPV16	Human papillomavirus type 16
HRP	Horseradish peroxidase
HSIL	High-grade squamous intraepithelial lesion
IC	Immunochromatograph
IF	Immunofluorescence
IgG1	Immunoglobulin type G1
IPTG	Isopropyl-β-D- thiogalactoside
KCl	Potassium chloride
KH_2PO_4	Potassium dihydrogen phosphate
LB	Lysogeny broth
LBC	Liquid-based cytology
LFA	Lateral flow assay
LoB	Limit of blank
LoD	Limit of detection

LOOCV	Leave-one-out cross validation
LSIL	Low-grade squamous intraepithelial lesion
LSPR	Localised surface plasmon resonance
mAb	Monoclonal antibody
MBA	4-mercaptobenzoic acid
MBN	4-mercaptobenzonitrile
MES	2-(N-morpholino)ethanesulfonic acid
MetOH	Methanol
MGITC	Malachite Green isothiocyanate
MP	4-mercaptophenol
MRI	Magnetic resonance imaging
NA	Numerical aperture
Na_2HPO_4	Disodium hydrogen phosphate
NaCl	Sodium chloride
NaF	Sodium fluoride
$NaHCO_3$	Sodium bicarbonate
NGS	Next-generation sequencing
NH_2	Amino group
NHS	N-hydroxysuccinimide
Ni^{+2}	Nickel ion
NP	Nanoparticle
NP-40	Nonyl phenoxypolyethoxylethanol
O-P-O	Phosphate backbone of DNA/RNA
PBS	Phosphate-buffered saline
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFA	Paraformaldehyde
pfu	Plaque-forming units
pH	Potential of hydrogen
pI	Isoelectric point
PIM	Protease inhibitor mix
PoC	Point of care
PPY	4-(1H-pyrazol-4-yl)pyridine
Protein G′ (PrG′)	Recombinant Protein G
PYOT	5-(pyridine-4-yl)-1,3,4-oxadiazole-2-thiol

Q	Quality factor
RIA	Radio immunoassay
RNA	Ribonucleic acid
R-NP	Reporter nanoparticle (also BPE-NP, etc.)
RPMI-1640	Roswell Park Memorial Institute 1640 Media
RT	Room temperature
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SERS	Surface-enhanced Raman spectroscopy
SiO_2	Silicon dioxide
siRNA	Small interfering ribonucleic acid
SNR	Signal-to-noise ratio
SPP	Surface plasmon polariton
SRS	Stimulated Raman spectroscopy
TAE	Tris-acetate-EDTA
TBS	Tris-buffered saline
TBT-T	TBS-Tween
TEM	Transmission electron microscope
TEMED	Tetramethylethylenediamine
Tet	Tetracycline
Tris	Tris(hydroxymethyl)aminomethane
TRITC	Tetramethylrhodamine
UV	Ultra violet
w/t	Wildtype
WMRS	Wavelength-modulated Raman spectroscopy
Xgal	5-Bromo-4-chloro-3-indolyl-β-D-galactoside

Dedicated to the memory of Jean Woolford, whose enthusiasm for her granddaughter's adventures in scientific research never wavered, but who never got to see me complete this thesis. You always said I got my brains from you - I think you might have been right.

Chapter 1

Introduction

Today, medical imaging plays a hugely significant role in the diagnosis and treatment of diseases such as cancer, from initial radiography and brightfield analysis of tumour margins to fluorescence-guided surgery and follow-up MRI scans. Relatively novel technologies have revolutionised disease management pathways and become a common sight in hospitals. Cervical cancer was one of the first cancers to receive the benefit of a diagnostic test, through the pioneering work of cytopathologist George Papanicolaou in the 1920s [1]. Its widespread clinical acceptance in the US and western Europe by the 1960s has significantly reduced mortality rates [2]. But despite general improvements to the technique, fundamentals remain unchanged, thereby failing to take advantage of the staggering advances in medicine, biochemistry and imaging since its inception.

One of those advances is the emergence of Raman spectroscopy, the measurement of inelastically scattered light from molecules, as a technique for sensitive diagnosis. Although the optical phenomenon was discovered by Chandrashekhara Venkata Raman in 1928, it did not reach its potential as an analytical technique until the advent of reliable laser light production in the 1960s. Raman spectroscopy, which is described in detail below, has since become an established technique in analytical chemistry applications and has evolved into multiple subspecies with often complementary strengths and weaknesses.

This thesis aims to consider which manifestation of Raman spectroscopy is best suited, in its current state, for improvements to the sensitivity and specificity of cervical screening. Once selected, that method will be developed and characterised in the light of relevant clinical factors to provide the framework for a new approach to screening. Owing to the variety of approaches and techniques used throughout the studies presented and for the reader's convenience, each Chapter (with the exception of General Materials and Methods) represents a contained body of work with its own introduction, methods, results, analysis and conclusion. This Chapter outlines the

overall research landscape, including the initial position of Raman spectroscopy as a tool for screening and the clinical problem which the proposed Raman spectroscopic studies seek to address.

1.1 Cervical cancer

An understanding of the epidemiology, mechanisms and current clinical pathways associated with cervical cancer and its precursor lesions is critical to the development of a relevant tool. Although the burden of disease is considered from global as well as local perspectives, reviews of current clinical practice are limited to those provided by the UK National Health Service. A recent review by Vale et al. considers screening practice in low and middle income countries [3].

1.1.1 Burden of disease

Cervical cancer is a treatable and preventable disease which remains, for women, the most common cancer in 38 countries and the leading cause of death in 45. It is the fourth most common cancer globally, accounting for 7.5% of all female cancer deaths [4]. Cancer and precancers of the cervix are particularly associated with younger women and particularly those living in economically underdeveloped areas. The disease disproportionately affects women of working age and so was ranked in 2013 amongst the top three causes of lost years of healthy life (Disability-Adjusted Life Years, DALYS) [5].

Another significant characteristic of cervical cancer epidemiology is the striking discrepancy in morbidity and mortality between low- and high-income regions on both a local and national scale. An 18-fold variation in morbidity has been documented between low and high risk regions in 2017, with 85% of new cases and 87% of deaths occurring in less developed regions, particularly in Africa [6]. And yet, only 5% of global cancer spending occurs in these areas. Figure 1.1 demonstrates cervical carcinoma incidence in the context of screening programmes.

Low and middle Human Development Index (HDI) countries are witnessing a shift in disease burden as non-communicable diseases overtake infectious disease and maternal complications as the primary causes of mortality. As a cancer overwhelmingly caused by prolonged infection with human papillomavirus (HPV), cervical cancer straddles these epidemiological categories, making effective population screening and affordable treatment critical in the absence of vaccination programmes [9]. In countries such as Turkey where screening has only recently been introduced, uptake is

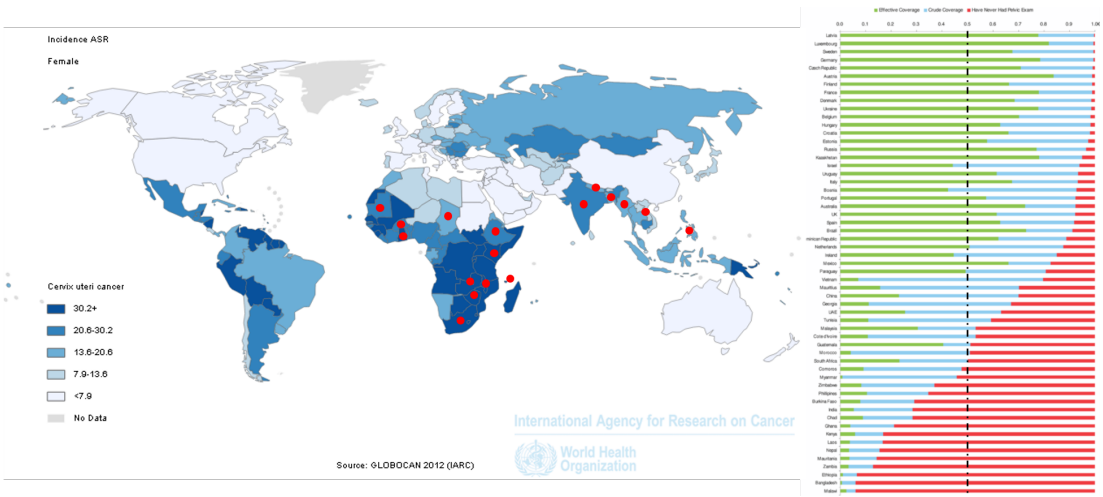


FIGURE 1.1: Left: Global incidence of cervical cancer in 2012 [7]. Right: Proportion of women, by country, attending regular screening (green), occasional screening (blue) and no screening (red) [8]. Countries with greater than 50% unscreened population are marked in red on the incidence map.

too low (1-2%) to be effective [10]. Other risk factors for cervical cancer and pre-cancer include smoking, prolonged use of oral contraceptives, and a weakened immune system (especially due to HIV infection) [11].

In Scotland, the overall incidence rate of cervical cancer has decreased by 6.3 diagnoses per 100,000 from 1981 to 2016, but incidence increased marginally from 2013-2016 [12]. This trend is reflected in other high-income countries and may be caused by decreased screening uptake, sexual behavioural changes and higher rates of HPV infection. Both screening and vaccination uptake are influenced by age and Scottish Index of Multiple Deprivation (SIMD) score, with younger women from poorer backgrounds demonstrating both lowest preventative uptake and highest HPV risk.

Proven approaches to reducing these inequalities include vaccination against HPV, population screening, surgery and medicines, with screening identified by the World Health Organisation as a 'Best Buy' for this disease. As a result of the effective implementation of these approaches in high HDI regions, the incidence rate of cervical cancer 1990-2013 decreased by 59%, leading to an increase in overall incidence of just 8.5% following population adjustment. Cervical cancer comprises up to a third of all cancers diagnosed (in both sexes) in many low HDI settings, compared with less than 10% in most very high HDI countries [9] and rates are three times lower in countries with screening programmes than without [5].

1.1.2 Mechanisms of disease

The pathway of HPV infection and subsequent cancer progression is reasonably well understood and is intrinsically linked to the process of differentiation in squamous epithelium [13]. HPV DNA is circular, double-stranded [14] and contains approximately 7900 base pairs [15]. It is coated with L1 and L2 capsid proteins and codes for proteins E1, E2, E4, E5, the oncogenes E6 and E7, and L1 and L2 themselves. The full sequence can be seen in Figure 1.2.

In normal epithelium, the basal layer produces cells which exit the cell cycle (1.2 and begin the process of terminal differentiation. The normal biochemical and morphological changes which occur as cells approach the surface of the epithelium are disrupted following HPV infection.

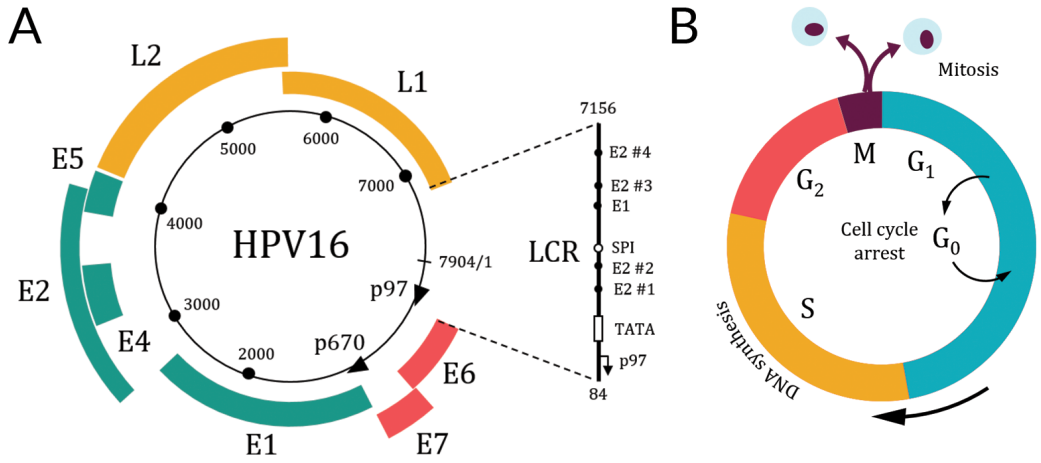


FIGURE 1.2: A) The genome of HPV16, with the long control region (LCR) expanded to show the E1 and E2 binding sites and TATA box. L1 and L2 are capsid proteins. E1,2,4 and 5 control viral expression and replication. E6 and E7 are viral oncogenic proteins which disrupt the cell cycle [16]. B)Stages of the cell cycle. G_1 = Gap 1 (readying for synthesis), S = synthesis of DNA, G_2 = Gap 2 (readying for mitosis), M = mitosis. G_0 (Gap 0) represents exit from the cell cycle.

The majority of HPV infections occur at the 'transformation zone' [17], where the columnar epithelium of the endocervix meets the squamous epithelium of the ectocervix [15]. The virus cannot infect terminally differentiated cells directly, but attaches via the L1 capsid protein to cells in the basal layer of the epithelium which are exposed following epithelial trauma [13]. L2 allows transfer of viral DNA into the basal cell nuclei, where it initially exists as an extrachromosomal element [16]. E1 and E2

proteins are then encoded, which are responsible for maintenance and replication of the virus in infected cells. E2 also controls expression from the viral early promoter (p97 in HPV-16), as well as of E6 and E7.

Active transcription of oncoproteins E6 and E7 then occurs. E6 binds to and degrades p53; a tumour suppressor protein that governs cell cycle arrest, apoptosis and DNA repair. E7 binds to retinoblastoma protein pRB, preventing it from forming a complex with E2F-1. Freed E2F-1 allows entrance of the cell into the S-phase of its cycle. Overall, this has the effect of forcing normally out-of-cycle cells to proliferate without usual checks for DNA integrity [15]: the full pathway and subsequent protein expression can be seen in Figure 1.3. E5 then increases cell proliferation and delays differentiation through increased protein kinase activity. Increased expression of E1,E2 and E4 drive genome amplification. Finally, the capsid proteins L1 and L2 are produced to package the new viral particles in cells. The completed virions are released as the outer layers of the epithelium shed. E4 is also expressed at this stage: its exact role is unclear, but it interacts with cytoskeletal proteins and is known to be involved in the production and release of virions [16]. The viral replication process is described graphically in Figure 1.4.

Critically, in normal productive HPV infection, only a small number of cells close to the basal layer re-enter the cell cycle and proliferate in order to increase the number of infected cells. Progression to cancer occurs when strict regulation of E6 and E7 are lost and 'runaway' proliferation occurs. This results in increasing numbers of replicating cells appearing towards the surface of the epithelium, replacing the normal differentiated structure (Figure 1.5). The productive replication cycle of HPV is then aborted. These abnormal HPV-driven cells expand to fill the epithelium, producing a high-grade squamous intraepithelial lesion (HSIL, equivalent to CIN 2/3), which may progress to invasive carcinoma if the basement membrane is breached. The most common metastatic sites are the lungs, liver, kidneys and lower intestinal tract.

It also appears that as cervical precancers progress, an increasing number of cells demonstrate integration of HPV DNA into the host cell genome [16]. The process through which this occurs is not fully understood as the integration sites vary in both number and location from cell to cell[18]. Integration is typically accompanied by loss or disruption of the coding sequences for E1, E2 and E4 and subsequent overexpression of E6 and E7 [16].

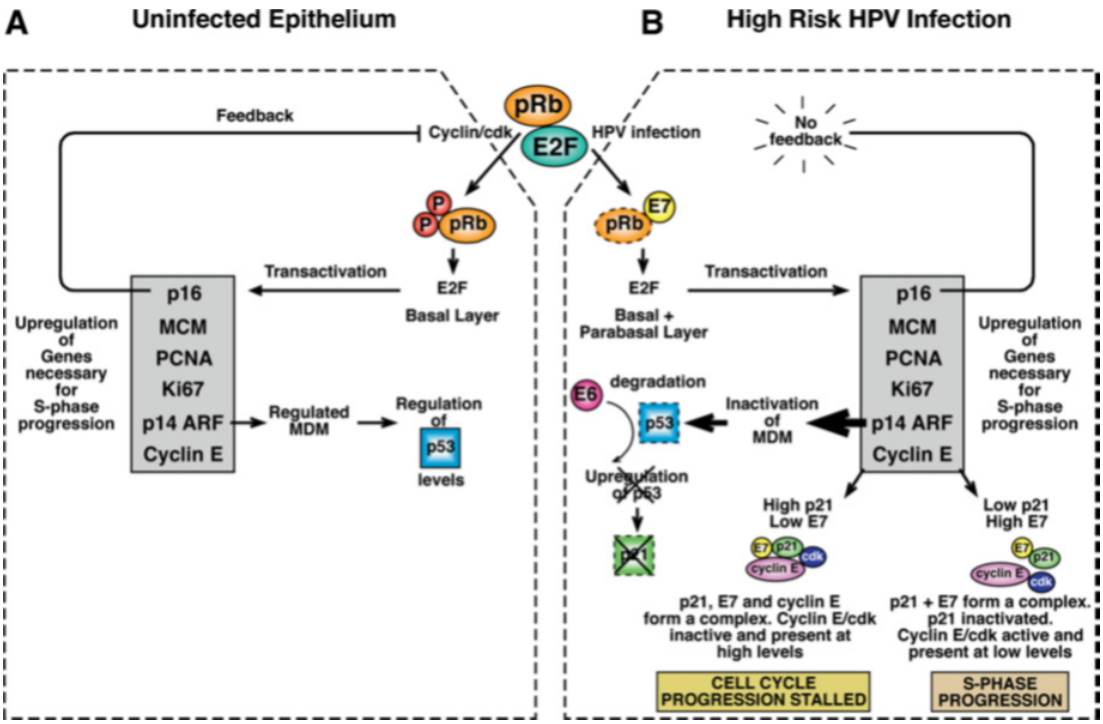


FIGURE 1.3: The cell-cycle deregulation pathway for high-risk HPV types. E6 and E7 bind to p53 and pRB respectively, disrupting their role in arrest and regulation of cell proliferation. In infected epithelium (right), this results in the up-regulation of proteins such as p16. HPV infection also appears to be dependent on p21 levels relative to E7 [16].

1.1.3 Clinical presentation of cervical cancer

Despite the well-defined progression from pre-invasive lesions to cervical carcinoma, patients are often asymptomatic until late stage disease. Abnormal vaginal bleeding is the most common symptom but is a non-specific indicator, occurring in several other gynaecological pathologies. Thus, screening is critical to identify disease as early as possible, with minimum disruption to the patient's quality of life.

If the patient is referred to colposcopy, the cervix is stained with acetic acid to highlight dysplastic areas. Biopsies are taken if deemed necessary. Tissue sections are H&E stained for analysis - typical examples of low- (LSIL) and high-grade squamous intraepithelial lesions (HSIL) and associated cytologies can be seen in Figure 1.6.

1.1.4 Current clinical practice

Current clinical practice varies geographically, reflecting differences in risk, population size, cost effectiveness and prevention strategies. The majority of high-income

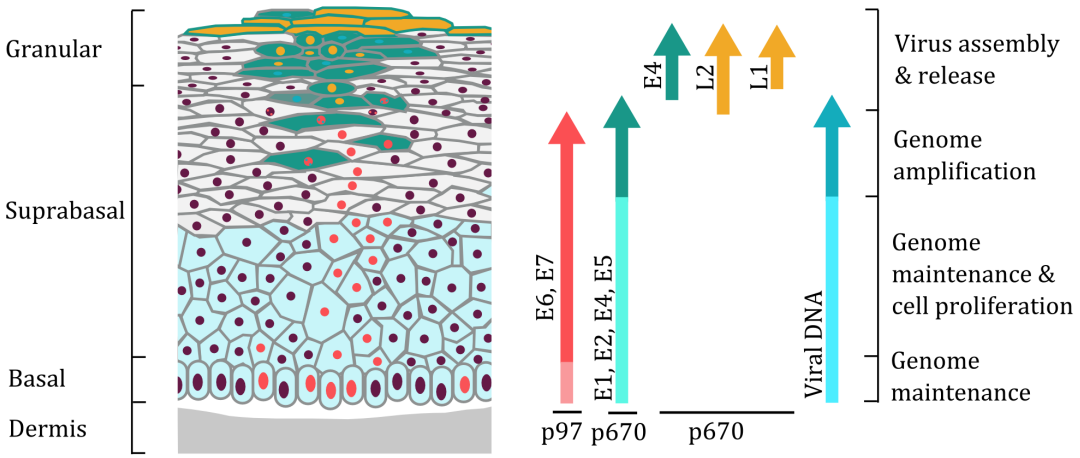


FIGURE 1.4: The HPV replication process with associated gene expression levels shown on the right. Greater intensity of colour corresponds to higher expression. Light blue and light grey cells are normal undifferentiated and differentiating cells respectively. Modified from J.Doorbar 2006 [16].

countries have established programmes of vaccination, screening and disease treatment [8].

The bivalent vaccine, providing specific protection against HPV16 and 18 and cross-protection against HPV31, 33, and 45, was introduced as part of a routine school-based immunisation programme in Scotland from 2008 to 2012, when it was replaced with the quadrivalent vaccine (HPV16, 18, 6 and 11). Data from the 1995-6 and later birth cohorts in the UK showed virtual eradication of HPV16 and 18 infection, with no significant evidence of replacement with other strains [20, 21].

The vaccine programme has achieved considerable success in Scotland and elsewhere, with the first vaccination cohort (1995-1996) demonstrating 89% reduction in prevalence of CIN3+ compared to the unvaccinated 1988 cohort. Herd protection was also demonstrated for non-vaccinated girls in the first cohort [22]. However vaccine uptake in this original cohort was high at 90%. A combination programme of vaccination and screening is still required. Cervical screening uptake amongst young women, and particularly unvaccinated young women, was low at 51 and 23% respectively [23]. Self-sampling HPV kits have been posited to increase likelihood of screening uptake [24].

The national screening programme is estimated to prevent 4,500 deaths in the UK every year [25]. However, such programmes require large, sustained financial input

and human resources, and are not completely accurate. Cytological sensitivity can be as low as 57% in some cases [26, 27], though specificity is typically considerably higher.

In a cytology screen, cells are brushed from the cervix and stored in fixing solution as a liquid-based cytology (LBC) sample, prepared according to the ThinPrep method [22]. This 'smear' is transferred to a microscope slide, stained and viewed by a cytopathologist, who classifies it as normal, borderline, low or high grade dyskaryosis, glandular neoplasia or unsatisfactory [22] (Figure 1.7). The interpretation of cell morphology is hugely complex and therefore open to subjectivity and false classification. In a 2008 clinical trial comparing reviewed with initial pre-treatment histologic diagnosis, the probability of unnecessary treatment between 8-27% was found (dependent on initial cytologic findings) [28], indicating that many women may be being referred to colposcopy unnecessarily. There is also evidence to suggest that there is a failure to detect cervical adenocarcinoma [29, 30], which makes up more than 10% of cervical carcinoma samples.

In the UK until recently, standard practice was to triage borderline cytology samples according to high-risk HPV status, although this is changing as HPV replaces cytology as the primary screen. In Scotland until recently, women were referred for colposcopy after one incidence of either high grade dyskaryosis or atypical squamous cells, two findings of low grade dyskaryosis, three findings of borderline changes after follow-up or three abnormal results in 10 years. Any low grade or borderline changes are followed with cytology after six months [22].

Current clinical practice will change in early 2020 in response to multiple studies which indicate that use of HPV genotyping instead of cytology as the primary cervical screen increases diagnostic sensitivity, thereby reducing the number of positive cases missed during screening [31, 32, 33, 34, 35]. Following a negative result for CIN2+, HPV-based screening also reduces the likelihood of incidence in the next five years, providing a case for increasing the screening interval and reducing costs [36, 37].

In the case of a high risk HPV-positive test, reflex liquid-based cytology (LBC) will be used to confirm the need for further tests. The new testing framework requires additional patient management streams - negative cytology results following HPV positivity will lead to increased monitoring and screening frequency. Solutions to this still vary globally [38]. In Australia, co-testing of HPV and LBC will be used for patients following treatment of high-grade abnormalities or who are deemed to be high-risk [39], whereas in the UK, colposcopy referral will be used [31].

Studies have shown that the move to HPV testing results in an increase in colposcopy referral of up to 80% [31], owing to the reduced specificity of HPV testing

compared with cytology. The requirement for retaining cytology to account for this reduction in specificity of the primary test indicates that a simultaneous test, or one that increases the overall accuracy of secondary cytology, would be beneficial to reducing the burden on colposcopy.

1.1.5 p16 as a marker of cervical cancer and pre-cancer

The use of biomarker screening has been posited as a potential method for improving classification accuracy in cytology, but identifying targets relevant to genetically heterogeneous populations exposed to a variety of environmental factors while maintaining high positive predictive value is often unfeasible [40]. However, the HPV-based origins of the vast majority of cervical cancers provide an opportunity for robust biomarker development which may not be available elsewhere.

Significant work is being undertaken to further understand cervical cancer and precancer pathology, and the biomarkers that arise during disease progression. A broadly applicable, robust test should be relevant to all HPV types and resistant to the changing pattern of disease. In the context of HPV-based primary screening, Bhatia et al. identify five requirements for an effective triage for hrHPV positive women: "high specificity for CIN2+; objectivity; high-throughput suitable for automation; amenable to a wide range of biospecimen types and produced independently of the infecting virus type." [41]

In a review by Herrington in 2001, two aspects of biomarker discovery were highlighted as areas for improvement [42]. The first was that diagnostic and prognostic markers often fail to distinguish between pathologies requiring different patient management strategies and are therefore not especially clinically useful. The second is that many of the studies performed have insufficient power to assess the value of the marker. Although HPV typing is highly sensitive for predicting lesions of grade CIN2 and higher, it is not capable of distinguishing between CIN grades overall and has low positive predictive value.

More recently, research has focused on the investigation of universal cellular proteins which are affected by HPV infection and disease progression, which constitute 'surrogate' markers. These include MCM5 [43] and CDC6 [44]. One of the most widely investigated markers is p16 protein expression, also known as $p16^{INK4A}$ or cyclin-dependent kinase inhibitor 2A (CDKN2A). In normal epithelium, p16 regulates expression of S-phase entry proteins (PCNA, MCM) by inhibiting CDK activation [45]. Following HPV infection, S-phase entry is determined by the E7 oncogene, which triggers overexpression of p16 in line with E7 levels as an overcompensation mechanism.

Critically, abnormally high cellular p16 expression is independent of the hrHPV type and so potentially obviates the need for type-specific testing.

Diffuse staining with p16 has been demonstrated for virtually all cervical carcinomas and high-grade lesions, as well as in CIN1 lesions which were shown to be at higher risk of progression [44]. Thus, it represents a useful clinical marker in terms of its ability to stratify lesions requiring treatment. Use of p16 as a biomarker for cervical disease has been considered in multiple contexts including stratification - primary screening, triage of screening results and cell-line cytology discrimination; diagnosis following colposcopy, and prognosis. This reflects the potential for use of p16 in different formats dependent on the patient management strategy chosen.

In a stratification context, it has been shown to be superior to conventional cytology in a trial of 1137 women when coupled with HPV testing [46]. A dual test for p16 and HPV E4 expression has been suggested for sample stratification [47, 48], but p16 is typically paired with Ki67. Several studies have demonstrated the utility of p16/Ki67 dual staining in enhancing cytology sensitivity [38]. Conflicting data has arisen elsewhere: no alteration in p16 expression was observed in ten cervical cell lines (8 HPV+ive, 2 HPV-ive) as a result of mutational inactivation of p16, although this is deemed to be a rare event [49] compared with many other studies indicating correlation of overexpression with increasing CIN grade.

Following colposcopy, staining of tissue sections with p16 has been shown to significantly improve interobserver agreement for diagnosis of CIN2, 3 and high-risk HPV CIN1 [50], and to be associated with all high-risk HPV infected lesions [43]. It has also been shown to be present in the majority of cases of invasive disease but not morbidities such as chronic cervicitis [51]. Conversely, a meta-analysis of p16 staining trials suggested that p16 staining positively associated with better prognosis in certain settings (no lymph node metastasis, disease-free survival) but did not significantly correlate with tumour size or grade [52].

Due to evidence of excess patients referred to colposcopy following both HPV- and cytology-based primary screening, it is considered that p16 staining, preferably with Ki67, is most beneficial as an adjunct to or replacement for cytology-based triage of hrHPV samples.

1.1.6 Alternative biomarkers for triage of hrHPV primary screening

Besides p16/Ki67 many classes of non HPV-specific biomarker are available, although none have been trialled in large-scale studies. These include proliferation and cell cycle control markers (p53, cyclin dependent kinases (CDKs)) and differentiation (CD44)

markers [53]. Other potential markers for cervical carcinoma include cancer stem cell markers such as SOX2 and TRA-1-60 [54] and host chemokine signatures including CCL2 [41], although such pro-inflammatory markers may be compounded by sample variability due to co-morbidities or hormonal changes.

HPV-specific markers of disease have been considered to a greater extent in the triage context. HPV genotyping provides another layer of information regarding the progression risk associated with a given type, but such information still has poor positive predictive value, which is compounded by the use of vaccination programmes which protect against the highest risk strains. The global variation in high-risk strains would also prevent implementation of a broadly applicable test.

Of greater promise are markers of self- and HPV-related markers of methylation, which have been shown to be directly correlated with cervical disease and do not require morphological study. Markers such as *MAL* and *CADM1* have been shown to provide sensitivities of 84% for CIN3 when considered simultaneously, approaching sensitivities previously found for p16/Ki67 classification of CIN3 [38]. It is not clear how many methylation markers would be required to constitute a robust test, and there have been no examples to date of direct comparisons between the accuracy of methylation and dual-staining based triage approaches.

With the appropriate application of technology to promote automated, potentially high-throughput classification as investigated in this thesis, p16 has the potential to fulfil all of the triage test requirements as outlined in Section 1.1.5. As a surrogate marker it provides a particular advantage over type-specific HPV testing. A direct comparison of dual p16/Ki67 expression with promising methods such as methylation assays which also considers cost and resource burdens (i.e. the possibility of self-taken samples, automated or non-specialist cytological analysis or triage of residual sample from the primary screen) would be of use. Whilst new combinations of markers may established as superior and adopted in the future, p16 is a robust, widespread and clinically useful marker in the screening process whether used alone or in conjunction with Ki67, and thus represents an ideal candidate for technological development.

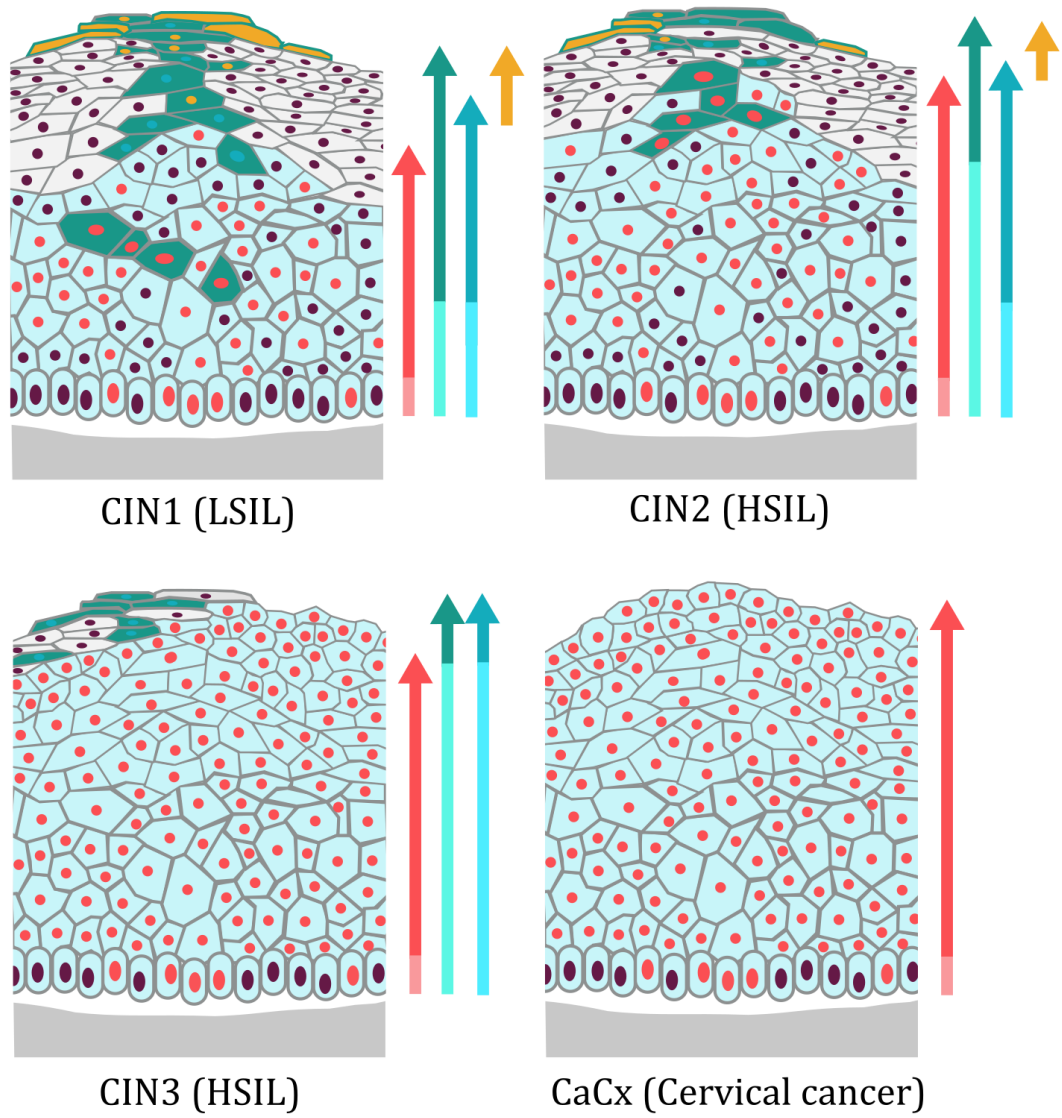


FIGURE 1.5: Progression from HPV infection to cervical cancer, characterised by an increase in proliferating cells (light blue) and eventual loss of HPV replication. The arrows and their relative colour intensity denote the expression levels of E6/7 (pink), E4 (green), viral DNA (blue) and L1 (yellow). The primary categorisation is that of Cervical Intraepithelial Neoplasia (CIN1,2,3), and the bracketed value is of the Bethesda system (High/low-grade Squamous Intraepithelial Lesion, H/LSIL). Modified from J.Doorbar 2006 [16].

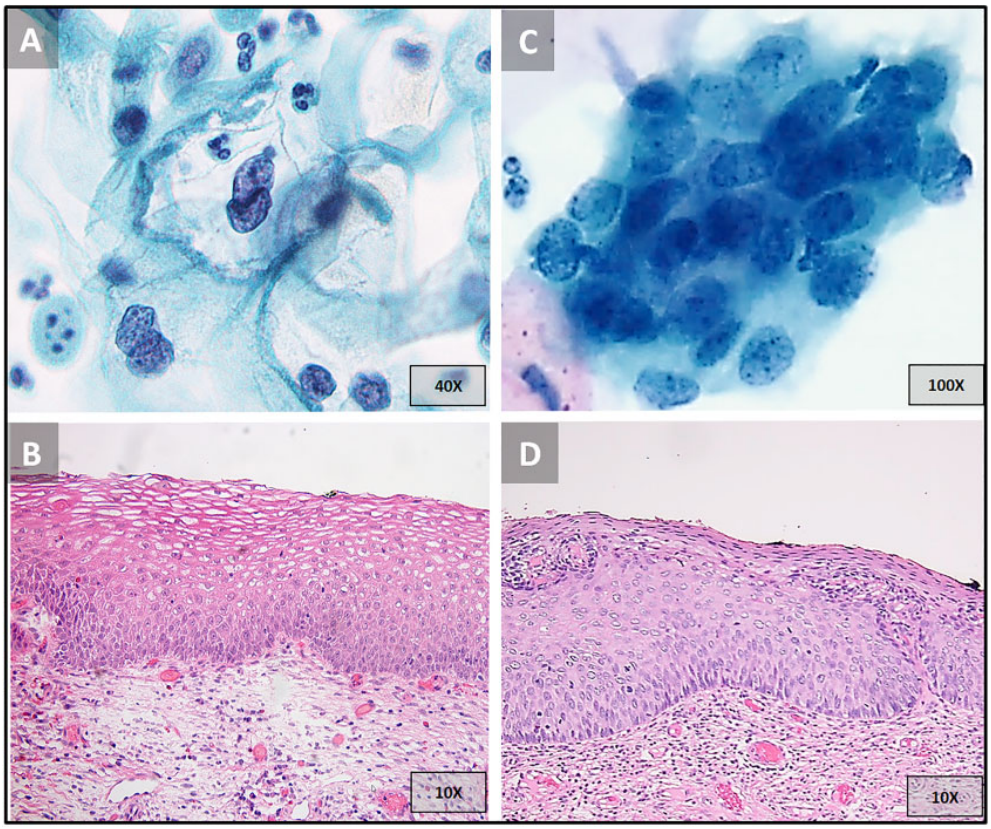


FIGURE 1.6: Cytologic and histologic images of low-grade SIL (A, B) and high-grade SIL (C, D). SIL indicates squamous intraepithelial lesion. Reproduced from Kim et al. 2018 [19].

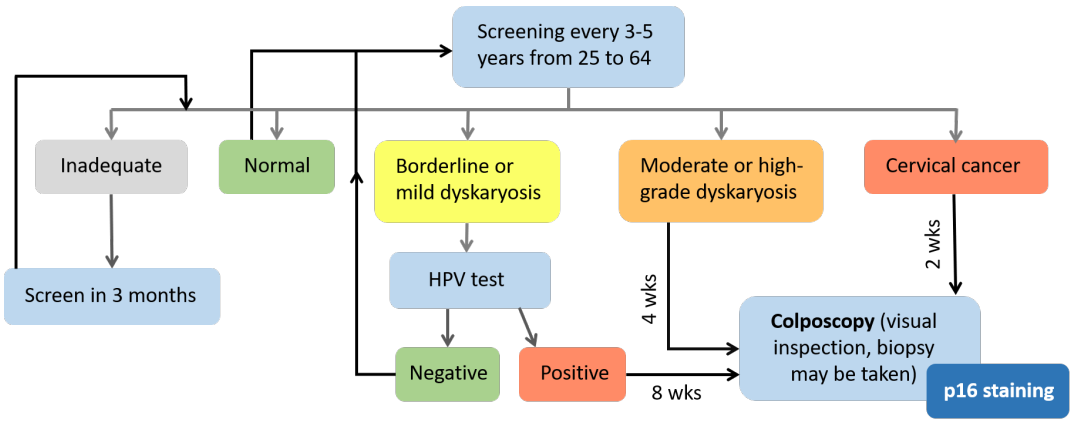


FIGURE 1.7: Former UK clinical pathway used for cervical screening. Dyskaryosis is typically defined by an increased nuclear to cytoplasmic ratio and abnormal chromatin pattern.

1.2 Molecular pathology

The use of biomarkers such as p16 or HPV DNA in diagnosis can be described as molecular pathology. In 2018 Coleman described molecular pathology as the application of the principles of basic molecular biology to the investigation of human disease processes [55]. It shares aspects of both anatomical pathology and molecular biology, and so is fundamentally interdisciplinary in nature. Differing pathologies and visual phenotypes typically used for diagnosis in this field are driven at a submicroscopic level by cellular machinery responding to genetic and environmental cues. Through consideration of the molecular content of cells, tissues and bodily fluids, a diagnosis can be made on the basis of genuine drivers or markers of disease rather than visual phenotypes which may be conflated by the most skilled pathologist.

Molecular pathology can be conducted alone or in combination with traditional histopathology or cytology approaches (morphomolecular pathology). Far from sounding the death knell for classical pathology, the knowledge gained from established approaches will be critical in ensuring that molecular diagnostics are suitably robust for clinical use [56, 57, 58]. Three challenges have been identified for the modern pathologist: the assertion of traditional pathology as a now-enriched practice; the integration of new molecular technologies into existing workflows; and the dissemination of molecular pathology knowledge amongst training and practising pathologists [59].

1.2.1 Global and targeted approaches

The different approaches to molecular pathology can be broadly grouped into two categories: 'global' or signature-based methods, and target-specific methods. Briefly, the signature-based approach relies on unspecified molecular differences between disease-positive and -negative samples, whereas the targeted approach chooses a likely biological marker for disease and quantifies target levels in the same samples.

Established examples of global approaches to molecular pathology include next-generation sequencing (NGS) and mass spectroscopy, and tend to act within a 'systems biology' framework, comparing populations of relevant molecules in the genome, transcriptome, proteome or metabolome [60]. These may represent either preliminary biomarker discovery processes to elucidate relevant targets, or methods of analysis in their own right. Mass spectrometry has been used to identify 55 dysregulated proteins [61] as well as potential metabolomic biomarkers for cervical cancer from plasma [62] and tissue lysate [63]. Thus far, NGS has only considered cervical cancer from an HPV perspective and therefore cannot be considered truly global in nature. It has been used

to consider HPV testing for self-sampled cytology [64] and in comparison with PCR and other methods [65].

Targeted approaches are currently the most familiar methods in the molecular pathologist's toolkit, though that may change as whole-sample sequencing and subsequent global bioinformatic analysis becomes steadily more accessible to non-specialist research groups. Immunohistochemistry (IHC) and immunofluorescence (IF) multiplexing is typically limited by the number of antibody host species available (unless directly labelled antibodies are made) but their use in biomedical research is almost ubiquitous. Hybrid capture and PCR or isothermal amplification-based methods can be used to detect the presence of HPV types by the highly conserved L1 gene [66]. This has been combined with liquid biopsy but with low detection rates (61.6%) [67]. Fluorescence in-situ hybridisation (FISH) allows for specific DNA or RNA strands to be localised. Other established targeted molecular pathology techniques are too numerous to mention here but are described by Best et al. 2016 [68].

1.2.2 Advantages and disadvantages of molecular pathology

Some of the advantages and disadvantages of using molecular pathology apply generally, but many are specific to the global or targeted contexts. Overall, molecular testing methods give the advantage of typically offering higher specificity, results independent of poorly preserved or altered morphology, and a biochemical signature for disease which may predate what can be detected through morphologial changes. The latter is particularly critical for reliable early-stage diagnosis which will not only reduce patient burden but also cost of treatment.

Global methods allow the full complexity of a given disease to be considered, which is useful from a clinical perspective when multiple mechanisms of disease result in the same treatment pathway. They also allow for unsupervised analysis of sample '-omics', thereby circumventing the need for time-consuming biomarker discovery. However, these methods typically require more expensive and time-consuming analysis. Targeted methods are typically cheaper and easier to analyse, which may have greater applicability in point-of-care or low resource settings.

As mentioned above, the use of molecular pathology alone does result in the loss of spatial data from traditional pathology, and so may have to be used as an adjunct to established methods especially in designation of tumour margins. There is suggestion that global methods such as NGS have resulted in 'information overflow' [69] which does not necessarily result in clinically useful information. As these methods often

inform biomarker discovery, they lead to potential issues with the specificity and biological relevance of targeted molecular techniques - if a single indicator is used, the diagnostic test may well omit patient populations. Too many indicators can lead to issues over cost effectiveness.

As will become clear, the Raman spectroscopy-specific approaches to global and targeted molecular pathology are largely shared with their non-optical counterparts. Global approaches are typified by data richness which is slow to acquire and analyse but capable of classifying multiple disease types. Targeted approaches are complementary - analysis and acquisition are much more rapid due to streamlining of data collection, but rely upon a robust marker or small panel of markers.

1.2.3 Current examples in clinical practice

Molecular pathology is now commonplace in some areas of clinical practice. Examples in oncology include ERBB2 (HER2) gene amplification in breast cancer [70], and mutation of EGFR, BRAF and BCR/ABL1 in lung cancer, melanoma and chronic myeloid leukaemia respectively [71].

Viruses [72, 73], bacterial infections [74] and sexually transmitted diseases [75, 76] have all been targeted for molecular diagnostic tests. There are currently four FDA-approved HPV tests available for cervical cytology samples which use PCR DNA amplification of HPV L1 (Roche Cobas), hybrid capture of whole HPV genome (Qiagen Digene), hybridisation protection assay of HPV E6/E7 mRNA (Hologic Gen-Probe) and unspecified DNA extraction (Hologic Cervista) [76].

A dual staining kit for p16 and Ki67 has been released by Roche (CINtec Plus Kit), with clinical utility confirmed in several studies [77, 78, 79, 80, 81]. It is designed for use in conjunction with the Cobas HPV test, ideally for samples which are HPV positive for non-HPV16 or 18 types. The patient management strategy for hrHPV primary testing is described in Figure 1.8. Although more effective than standard Pap cytology, staining protocols are still time-consuming and require laboratory apparatus and pathology expertise not available in areas of high unmet clinical need.

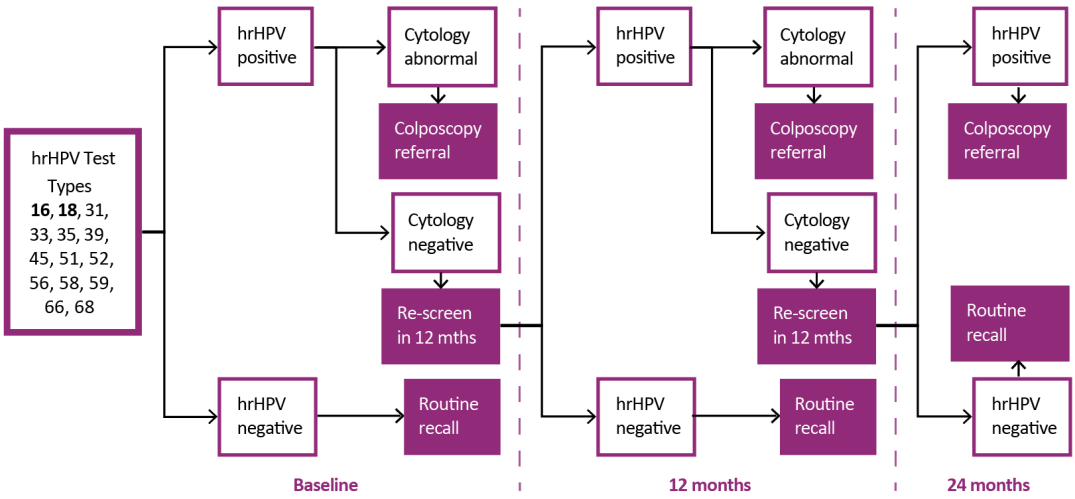


FIGURE 1.8: The patient management strategy for cervical cancer through molecular pathology methods, introduced in England by the end of 2019 and Scotland in April 2020. Cytology is aided by p16/Ki67 dual-staining. Adapted from resources by Public Health England [82].

1.3 Raman spectroscopy as a tool for molecular pathology

It is clear that a screening method able to employ quantitative molecular pathology in an automated fashion would greatly improve diagnosis, as well as reducing the cost burden through superior patient management and reduced resource requirements. Based on the inelastic scattering of light from molecular bonds, Raman spectroscopy is ideally placed to offer a more objective measurement of cervical cancer progression.

1.3.1 Fundamentals of Raman spectroscopy

Raman scattering is the inelastic counterpart to the Rayleigh scattering of light by molecules. In spontaneous Raman scattering, a photon with a given wavelength $\lambda_{incident}$ interacts with a molecule in a real energy state. The molecule is promoted to a higher vibrational energy and emits a photon in order to relax back to a real energy state. Elastic (Rayleigh) scattering occurs when $\lambda_{incident} = \lambda_{scattered}$. Stokes and anti-Stokes photons are emitted if the molecule relaxes to a higher or lower vibrational energy respectively (Figure 1.9). The shift in frequency of these Raman scattered photons indicates the energy difference between the intrinsic vibrational states of the system. This allows vibrational bonds of all the molecules in the laser focus to be characterised simultaneously, and broadly grouped into biomolecular categories (Figure 1.9).

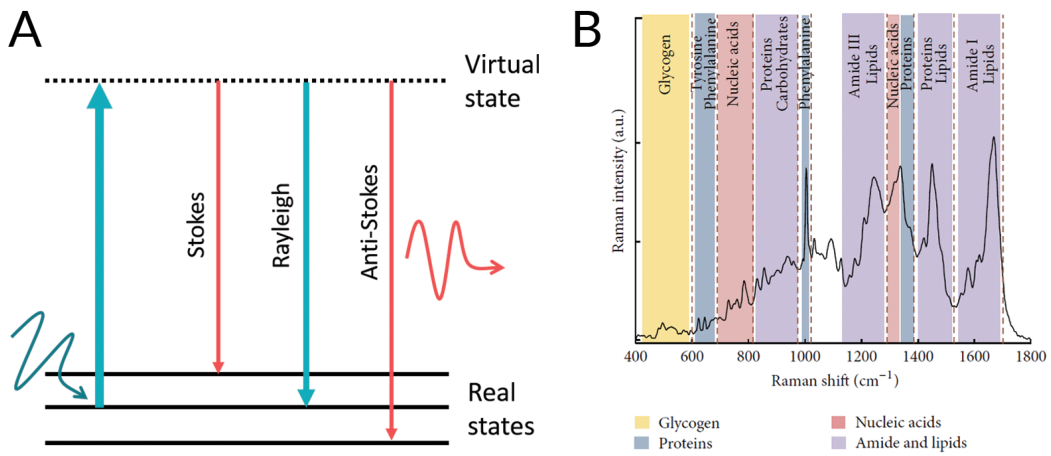


FIGURE 1.9: A) The energy level diagram for Rayleigh and spontaneous, non-resonant Raman scattering. Light of energy $h\nu$ interacts with a molecule in a real vibrational state. This promotes the molecule to a higher energy state, which typically is virtual. B) Raman spectral peaks can be broadly grouped into biomolecule categories based on unique vibrational modes occuring within these moieties [17].

Classically, spontaneous Raman scattering can be described by an electric field E which induces a time-dependent polarisability in a molecule. Undisturbed, the permanent dipole moment of a diatomic molecule oscillates at its natural frequency as the atoms vibrate. If the molecule is subject to an incoming photon with an oscillating electric field $\vec{E}(\vec{r},t) = \vec{E}_i^0 \cos(\vec{k}_i \cdot \vec{r} - \omega_i t)$, a dipole moment is induced, as the nuclei are attracted to the negative pole of the field, and the electrons to the positive [83]. Here $\vec{E}_i^0$ is the electric field amplitude, which includes the polarisation of the field. The incident wavevector and angular frequency are $\vec{k}_i$ and ω_i respectively. The induced dipole is then

$$\vec{\mu}_{induced} = \alpha \vec{E} = \alpha \vec{E}_i^0 \cos(\vec{k}_i \cdot \vec{r} - \omega_i t) \tag{1.1}$$

where α is the polarisability of the molecule. The magnitude of α varies as the molecule oscillates, and so can be written as

$$\alpha = \alpha_0 + (\Delta\alpha)\cos(\vec{q} \cdot \vec{r} - \omega_0 t) \tag{1.2}$$

where $\Delta\alpha$ is the maximum variation and $\vec{q}$ is the vibration (phonon) wavevector. Substituting Equation (1.2) into Equation (1.1),

$$\mu_{induced} = \alpha_0 \vec{E}_i^0 \cos(\vec{k}_i \cdot \vec{r} - \omega_i t) + (\Delta\alpha)\vec{E}_i^0 \cos(\vec{q} \cdot \vec{r} - \omega_0 t)\cos(\vec{k}_i \cdot \vec{r} - \omega_i t). \tag{1.3}$$

Using the trigonometric identity $\cos(\theta)\cos(\phi) = \frac{1}{2}[\cos(\theta+\phi)+\cos(\theta-\phi)]$ and collecting terms, the induced dipole moment is thus

$$\mu_{induced} = \alpha_0 \vec{E}_i^0 \cos(\vec{k}_i \cdot \vec{r} - \omega_i t) \tag{1.4}$$

$$+ \frac{1}{2}(\Delta\alpha)\vec{E}_i^0[\cos((\vec{k}_i - \vec{q}) \cdot \vec{r} - (\omega_0 - \omega_i)t) \tag{1.5}$$

$$+ \cos((\vec{k}_i + \vec{q}) \cdot \vec{r} - (\omega_0 + \omega_i)t)]. \tag{1.6}$$

Expression 1.4 is the Rayleigh scattering term, where the dipole oscillates and emits at the frequency ω_i of the incoming wave. Expressions 1.5 and 1.6 are the Raman scattered Stokes and anti-Stokes photons, which have a lower ($\omega_i - \omega_0 = \omega_f$) and higher ($\omega_i + \omega_0 = \omega_f$) energy than the incident photon respectively. The incident wavelength can be scattered by any molecule within the focus of the laser, meaning that information from all molecular species can be collected using the same laser frequency. The rarity of spontaneously occurring inelastic transitions results in a weak signal. This is especially the case for biomolecules, whose small Raman cross-sections (10^{-29}cm^2) result in just one in every 10 million incident photons undergoing Raman scattering [84]. The small number of Raman scattered photons means that long acquisition times are required to reach a reasonable signal to noise ratio (SNR).

Surface-enhanced Raman spectroscopy (SERS) overcomes the problem of inherently low Raman signal through amplification of the electric fields of incident and emitted light. This enhancement is provided by placing the molecule of interest in the vicinity of a metal surface, which is subject to the same laser source as in conventional Raman spectroscopy. The oscillating electric field of the laser photons produces a collective oscillation of conduction electrons (plasmons) in the metal surface. This is known as dipole localised surface plasmon resonance (LSPR). The incoming wavelength can be tuned to the resonant frequency of plasmons on a particular metal surface in order to maximise this effect [85], designated as surface-enhanced resonant Raman spectroscopy (SERRS). This localised electric field contributes to the intensity of both incident and Raman scattered light. LSPRs and the SERS effect are discussed in greater detail in Chapter 4.

Another modification to Raman spectroscopy to improve SNR involves reduction of non-vibrational background as opposed to amplification of the vibrational signal. There are multiple approaches to this which are described in Chapter 3, including wavelength-modulation of the incident laser light as first published by Shreve et al. in 1992 [86].

1.3.2 Raman spectroscopy in molecular pathology

Owing to the advantages of Raman spectroscopy as outlined above, the technique has become of increasing interest in biochemical analysis and medical imaging. It is capable of providing some of the specificity of established techniques such as immunofluorescence or mass spectroscopy but with a reduced need for sample preparation. The following section considers Raman spectroscopic progress in a general molecular pathology context, before cervical cancer screening is considered in more detail below.

Global approaches

A global Raman spectroscopy-based approach to molecular pathology is characterised by the use of intrinsic biochemical properties of the sample. Changes to the biochemical content of the sample, which accompany changes to disease state, are reflected in the Raman spectra. This 'intrinsic' measurement boasts the considerable advantages of being truly non-invasive and requiring negligible sample preparation. However, analysis of acquired data can be complicated. Fractional spectral differences between disease states, and the confounding influence of patient inter-variability, can expose the technique to overfitted data models and subsequent difficulties in the clinically relevant stratification of samples. The global approach typically makes use of 'standard' spontaneous Raman spectroscopy, or its label-free derivatives such as coherent anti-Stokes Raman spectroscopy (CARS) or stimulated Raman spectroscopy (SRS). SRS is considered in greater detail in Chapter 5.

The non-invasive nature of Raman spectroscopy and its ability to probe samples with vibrational specificity has resulted in an increasing interest in its use in a biomedical context, and particularly in the field of molecular oncology. Raman spectroscopic analysis has been employed extensively for the grading of both histology and cytology samples in a variety of cancer types [87], as well as for the determination of breast tumour margins in animal models [88]. Excellent reviews in this area are provided by Austin et al. [89], Krafft et al. [90], Kong et al. [91] and the Xie group [92]. In addition to cancer discrimination in tissue, Raman microspectroscopy has been used to consider cellular status, including for apoptotic and necrotic cells [93] and stem cell differentiation [94, 95]. Its label-free nature has provided significant progress in *in-vivo* imaging applications [96, 97].

Owing to the signal strength limitations of Raman spectroscopy, groups have sought to combine the technique with other global modalities. This has included optical

coherence tomography, mass spectrometry and somewhat surprisingly, autofluorescence [98].

Targeted approaches

Unlike the global approach, targeted Raman molecular pathology makes significant use of the understanding of specific biochemical changes within cells and tissues in order to diagnose disease. Disease markers, such as overexpression of a particular protein, are chosen for the target pathology and used to stratify disease state. This has the advantage of providing a more specific diagnosis, as levels of known markers can be quantified more easily than a 'blind' spectrum of vibrational changes. However, it does require significant sample preparation, as well as testing of appropriate markers and probes, in order to guarantee a true and repeatable diagnosis in clinical samples. The targeted approach is particularly suited to surface-enhanced Raman spectroscopy (SERS), which provides signal increases on the order of 10^{10} compared to spontaneous Raman spectroscopy [99]. Critically, this means that the acquisition time for a comparable SNR is greatly reduced.

In order to produce plasmonic field enhancement at levels practical for spectroscopic applications, the available surface area can be increased by the use of either rough surfaces or particles with nanoscale features. The improvement of nanoscale production and measurement techniques in recent years has seen an increased interest in the use of metal nanoparticles in solution with the analyte. This is known as 'conferred enhancement' of the SERS phenomenon. Both direct and indirect signal enhancement have been explored using nanoparticles. The former enhances the Raman signal of an endogenous compound in the sample and has been used to detect viral pathogens [100, 101, 102] and to image intracellular components [103, 104]. However, the enhanced Raman spectra in this case are as complex as their spontaneous predecessors. Reviews judge that the clinical translation of direct SERS methods will be extremely challenging, as clinical samples are characterised by both low ratios of target-to-irrelevant species and a high background [105, 106].

Indirect SERS involves generation of 'functionalised' nanoparticles containing an exogenous Raman label, a process which is described in detail in Chapter 4. Due to its ability to provide excellent acquisition times and a technique for quantifying marker levels which is similar to current immunocytochemistry methods, SERS has recently attracted a great deal of interest from the cell biology and disease diagnosis research communities. SERS also provides advantages over other mainstream targeted molecular methods, including fluorescence spectroscopy and immunological assays,

through sharper spectral features and greater detection sensitivity respectively. SERS has been used to discriminate lung, breast and leukaemia cell pathology, as well as glucose levels in animal models [99]. Studies by Chou et al. of the β-amyloid peptide found in Alzheimers-related plaques suggested that different secondary protein conformations could be detected using SERS with unfunctionalised gold nanoparticles [107]. The use of a nanofluidic device in this and other studies [108] for trapping and visualising nanoparticles and analytes is something that may be considered for automated diagnosis, alongside flow cytometry techniques. Endocytic uptake and internal transportation studies have taken advantage of the stealth properties of indirect gold nanoparticle-based imaging [109, 110].

Marker-specific SERS approaches have been employed in normal and cancer cell discrimination. Yang et al.s' study on epidermal growth factor receptor 2 (HER2) in breast cancer cell lines (SKBR3 and MCF7) used silver nanoparticles with p-mercaptobenzoic acid acting as Raman reporter and antibody conjugation agent [111]. This removes a step from the nanoparticle fabrication process. The sensitivity of SERS with antibody conjugated nanoparticles is such that Wang et al. were able to detect pancreatic cancer marker MUC4 in human sera at levels which could not be resolved by conventional formats such as ELISA [112]. This may prove to be a critical advantage for early detection methods, which could also be developed into a less time-consuming alternative to immunohistochemical techniques.

1.3.3 Raman spectroscopy in cervical cancer diagnosis

Global approaches

The first cervical cancer diagnostic investigations to employ Raman spectroscopy were executed by the Mahadevan-Jansen group in 1998, using both *in vivo* and *ex vivo* samples [113]. *In vivo* sampling was performed using a fibre-optic probe in the near infrared, indicating the primary interest in generating an endoscopy-like diagnostic tool. It was hypothesised that the ratio of peaks 1659:1321cm^{-1} and 1659:1452cm^{-1} could be used as an indicator of disease state. Although the probe was able to collect *in vivo* Raman data successfully, the unassigned peaks were used as a disease indicator instead of the full spectrum with multivariate analysis. This could potentially lead to misinformation due to inter-patient variability, which could not be verified due to only one abnormal and one normal patient being investigated. It must also be noted that as in other cervical sample investigations peak intensity differed between *in vivo* and *ex vivo* sampling. This suggests that analytical techniques developed for one sampling methodology are unlikely to be directly translatable. Further *in vivo* investigations of

tissue samples by the same group in 2001 suggested that peaks assigned to phospholipids and DNA undergo change with lesion progression [114]. However, difficulties were experienced with the heterogeneous nature of tissue, as well as spectral contributions from vibrational modes not specific to a particular pathology. This can lead to problems in training the algorithms used to analyse data, as a peak for one mode cannot be assigned uniquely to a marker of disease. The problem of tissue heterogeneity has been raised in reviews as a major barrier to the use of Raman spectroscopy as a 'digital histopathology' technique [87, 84].

The complexity of tissue signal analysis suggests that it would be beneficial to consider cytology samples, not simply as a precursor to *in vivo* Raman histopathology, but as a useful research outcome in its own right. This becomes even more relevant in the context of current cervical screening techniques. To date, five cell line [1, 115, 116, 117, 118] and two acquired cytology studies [1, 119] have been performed. The Jess et al. 2007 study [117] investigated the ability of confocal Raman spectroscopy to discriminate between the nuclei of primary human keratinocytes (PHKs), those expressing the E7 gene of HPV-16, and the cervical carcinoma-derived cell line CaSki. Up to 100% sensitivity was achieved for normal vs. carcinoma discrimination, although specificity values were marginally lower. Both live and 70% ethanol-fixed samples were investigated using a 785nm diode laser at 120s acquisition time per cell. The fixed cells demonstrated higher peak discrimination than the live cells, which is a helpful result for LBC sample consideration. The resulting spectra were analysed with the commonly used Principal Component Analysis (PCA) technique, which is described in Chapter 3. All peak differences for each cell type save 720, 1129 and 1452cm^{-1} were found to exceed the standard error and therefore be statistically (and diagnostically) significant. The paper also provides a full analysis of the individual vibrationally assigned peaks used in the PCA training algorithm as well as an explanation of their potential pathological significance. Such analysis lends greater specificity to the acquired data (as opposed to 'blind' analysis), which may potentially lead to a greater understanding between the imaging technique and underlying biochemistry of the disease. The main practical drawback of this experiment is the acquisition time required for each cell using standard Raman spectroscopy. A single abnormal cell in one cervical smear of 10,000 cells is sufficient to provide a positive diagnosis, meaning that a single slide would require approximately two weeks to image fully using point sampling.

Further cell line experiments by Ostrowska et al. investigated cervical cell recognition in relation to HPV type and viral copy number in C33A (HPV-negative), HeLa (HPV-18, 20-50 copies), SiHa (HPV-16, 1-2 copies) and CaSki cells (HPV-16, 60-600

copies) [115]. The results suggest that more significant spectral change (particularly of protein, DNA and lipid peaks) does correlate with increasing HPV copy number, but that PCA categorisation is more significantly affected by type. This highlights the potential of Raman spectroscopy to diagnose HPV type as well as cancer status, but does raise questions of how the diagnostic algorithm would approach a sample containing multiple infections. The analysis appears to be significantly affected by peaks assigned to Amide I and III, indicating a change in secondary protein structure. Increases in amino acids tryptophan, proline, tyrosine and phenylalanine could also constitute markers, as the same changes were seen in the 2007 Herrington group investigation for other cervical cell lines [117].

The move to clinical samples has only occurred for cervical cancer in recent years, with Rubina et al. publishing one of the first studies in 2013. 79% classification efficiency [119] was achieved. After spectral acquisition, the cells were stained and independently examined by two cytopathologists to give 45 normal and 49 abnormal specimens. An initial test of 17 normal and 20 abnormal samples considered using Principal Component-Linear Discriminant Analysis (PC-LDA) suggested that the main spectral differentiation could be attributed to the presence of blood components in the samples. Although blood contamination can accompany positive pathologies it cannot be exclusively attributed to disease. Thus, the remaining samples were subjected to red blood corpuscle (RBC) lysis buffer. The resulting difference spectra, calculated by subtracting the mean normal spectrum from mean abnormal spectrum, provides an initial indication of which peaks change with disease state. Although this differs from the cell-line spectral results of the Herrington group [117], the significant increase in Amide III seen in all of the above studies suggests that a more in-depth investigation of changes to secondary protein structure as a marker of cancer progression may be worthwhile from a global diagnosis perspective.

The potential problem of inter-patient variability has been investigated in two papers by the Mahadevan-Jansen group. The first focused on the effects of hormonal status on the biochemical fingerprint of cervical smears [120]. It found an increase from 88% to 94% in classification accuracy when diagnosis of normal, metaplasia, LGSIL, and HGSIL samples was limited to a single hormonal status (pre-menopausal). The second considered the effect of normal variation on disease classification in 172 patients[121]. These were categorised as 'true normal' (no history of disease); previous disease normal (disease history but presently normal); adjacent normal (spectral acquisition from normal area neighbouring diseased area); LSIL and HSIL. A Raman spectroscopic analysis of the effects of acetic acid used to highlight diseased areas

during colposcopy was also carried out. This was found to have no statistically significant repercussions for Raman analysis, which is a positive result for practical sample acquisition. Only four of the 146 classified spectra were misdiagnosed, providing a classification accuracy of 97%.

Further *in vivo* studies have yielded diagnostic accuracies of 82.9% using genetic algorithm-partial least squares-discriminant analysis and double cross-validation [122]. Higher diagnostic sensitivities were achieved by the same group using both low- and high-wavenumber spectral regions [123].

The first paper devoted to the use of SERS with cervical cancer was published in 2013, but used the enhancement method to improve spectral quality in a global approach. Here, Feng et al. investigated the use of blood serum samples from 60 cancer patients (38 stage 1-2, 22 stage 3-4) and 50 healthy volunteers for diagnosis of cervical cancer with direct (non-functionalised) SERS[124]. The prospect of using blood samples for diagnosis is appealing due to the ease of collection and transportation. PCA-LDA was used to discriminate between serum samples with 96.7% sensitivity and 92.0% specificity. These results are promising, but the major drawback of this approach is the specificity of non-functionalised SERS when applied to clinical screening. Diagnosis of cervical cancer using blood samples from the general population is likely to be confounded by other disease markers.

Targeted approaches

To date only one study considering p16-targeted SERS-active conjugates has been published [125]. Here, synthesised squaraine dyes were used as Raman labels on PEG-encapsulated gold NPs (40nm). Antibodies for epithelial growth factor receptor (EGFR) and p16/Ki67 were conjugated to these nanotags. HeLa cells were incubated with the nanotags and SERS spectra were recorded for live cells in PBS, with an acquisition time of 0.5s per cell. The nanotags for EGFR and p16/Ki67 bind to cell surface receptors and nuclear receptors respectively. This ability to introduce nanotags to the cell nucleus could be a significant advantage for the investigation of cervical cancer due to the importance of viral oncoprotein expression location. This has been shown to vary between high and low-risk HPV types [126]. As a proof of concept this investigation is a critical step in the use of SERS for marker quantification, but other cell lines and controls will have to be used in order to test its diagnostic capability.

More recent publications, specific to the studies undertaken, are reviewed in the introductions of Chapters 3 to 6. Overall, Raman spectroscopy has been shown to be a valid diagnostic tool, but one which is currently hindered by several factors. Overall,

there are a lack of sufficiently sized clinical trials to provide the necessary data for cost-benefit analysis and clinical translation. In the global case, slow acquisition times and sampling methods are barriers to progress. In targeted approaches, there appears to be a lack of collaborative research between basic scientists and research clinicians, resulting in well-characterised labelling systems which demonstrate novel optics or chemistry but have low clinical utility.

1.4 Exploring Raman spectroscopy as a tool for improved cervical screening practice

Cervical cancer screening in its current form is a successful programme but one which would greatly benefit from a molecular pathology approach in order to reduce workload and cost, increase objectivity and accuracy, and make screening more accessible to the demographics which need it most.

The fundamentals of Raman spectroscopy and previous studies in general biochemical and cervical carcinoma-specific frameworks demonstrate its potential to provide a sensitive and objective molecular pathology test, but one which requires the addition of enhanced Raman approaches to be clinically useful. This may take the form of a diagnostic capable of co-testing for HPV and cytological abnormalities, or one which replaces or provides objective guidance for current cytology practice.

The data presented above suggests that cervical cancer provides an ideal case study for the development of Raman-based spectroscopic tests, as a disease which requires greater diagnostic sensitivity in cytologic screening and also has the benefit of a robust surrogate marker. The decision to approach the project in terms of global and targeted approaches to diagnosis aligns with the current clinical reality and is therefore best placed to produce a practical test.

The aim of this thesis is to produce an objective, quantifiable and hopefully automated method of diagnosing cervical cancer and its early precancerous progression stages. This will use Raman spectroscopy as a form of molecular pathology analysis. In the first instance, the global (Chapter 3) and targeted approaches (Chapter 5) will be studied through the use of wavelength-modulated and surface-enhanced Raman spectroscopy respectively, following production of a functional SERS-active conjugate (Chapter 4). This overview data will provide the basis for selection of a clinically useful method for further development (Chapter 6).

Chapter 2

General materials and methods

General Materials and Methods comprises techniques which feature in more than one chapter. The majority of methods used in this project are specific to each aspect and so are detailed in the relevant chapter for ease of reference.

2.1 Cell culture

Operations involving cells prior to fixation were performed in a biosafety cabinet with laminar flow hood (TriMAT^2 Class II Microbiological Safety Cabinet) using standard aseptic procedures. Molecular biology grade reagents and plasticware were used, otherwise autoclaving was carried out. All maintenance and procedural incubation steps were at 37°C in 5% CO_2 using Cellstar culture vessels from Greiner Bio-One.

2.1.1 Primary cells

Primary cells and culture reagents were purchased from Lonza unless otherwise stated. The recommended culturing procedure for Normal Human Epithelial Keratinocytes (NHEKs, neonatal foreskin) resulted in reduced cell viability due to poor disassociation, cell clumping after centrifugation, and low cell adherence. The protocol was modified as described below to mitigate for this. NHEKs were subcultured for no more than 20 passages using ReagentPack Subculture Reagents from Lonza (Cat # CC-5034).

75cm ventilated culture flasks were washed once with 5ml 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) Buffered Saline Solution before adding 2ml 0.25% Trypsin/ethylenediaminetetraacetic acid (EDTA) at room temperature (RT). Plates were incubated at 1min intervals until 90% cell disassociation was confirmed with brightfield microscopy. Trypsinisation was neutralised with 4ml RT Trypsin Neutralising Solution and the cells counted using a 10µl sample prepared with 1:1 0.4%

Trypan Blue Stain (Sigma) and a Luna Automated Cell Counter. New plates were prepared with an appropriate volume of KGM-Gold medium at 37°C, and cells plated at the recommended seeding density of 10,000 cells/cm^2. KGM-Gold medium was supplemented with 0.1% hydrocortisone, 0.1% transferrin, 0.05% epinephrine, 0.1% gentamicin sulfate/amphotericin (GA-1000), 0.4% bovine pituitary extract, 0.1% recombinant human epidermal growth factor (rhEGF) and 0.1% insulin (all [v/v]) as provided with the media kit (Cat. # 00195769). 90% media replacement was carried out on alternating days.

2.1.2 Established cell lines

All procedures described are for 10cm culture plates using Gibco reagents unless otherwise stated. Cell stocks from established lines were maintained in the appropriate medium as described in Table 2.1, supplemented with 10% foetal bovine serum [v/v] (FBS) and 1% penicillin-streptomycin solution [v/v] ('Pen Strep' 10,000 U/ml, Thermo Scientific). All lines were subcultured 1-2 times per week as described below, dependent on growth rate and seeding density (1:3 to 1:10 dilution ratio).

Cell line	Culture medium	Trypsin /EDTA	Cell provenance	Provider
SiHa	DMEM	0.05%	HPV16 positive squamous cell carcinoma, adult cervix	ATCC
CaSki	RPMI-1640	0.25%	HPV16 positive squamous cell carcinoma, adult cervix (metastasis to small intestine)	ATCC
HaCaT	DMEM	0.25%	Immortalised keratinocytes, adult skin	Dr Duncan Sproul

TABLE 2.1: The cell lines cultured for all aspects of the project. See (1) for details of abbreviations.

Cell culture

For subculturing, culturing medium was removed and the cells washed with 2ml sterile phosphate-buffered saline (PBS) before 2ml RT Trypsin/EDTA at appropriate concentration was added (see Table 2.1) and the plates incubated for no more than 5min at

37°C. Dissociation was confirmed using brightfield microscopy before 8ml appropriate warmed medium was added. The resulting cell suspension was diluted into prepared culture plates containing the required volume of warmed medium. The plates were swirled to promote seeding uniformity and maintained for at least 24h before further treatment.

Cell freezing and recovery

For cell storage, confluent 10cm plates were incubated at 37°C for 5min with 1ml Trypsin/EDTA after washing with 2ml PBS to remove residual medium. Disassociated cells were mixed with 4ml appropriate serum-free media, transferred to a 15ml Falcon tube and centrifuged at 200g for 5min at RT. The supernatant was discarded, replaced with 3ml freezing media (50% FBS, 20% dimethyl sulfoxide (DMSO, Sigma), 30% serum-free media [v/v]) and the suspension aliquoted into 3 labelled CryoTube Vials (Thermo Scientific). The vials were transferred to a prepared freezing container and stored at -80°C overnight before long-term storage in liquid nitrogen.

For recovery, culture plates were prepared with 8ml appropriate medium and incubated at 37°C. Cryo vials were thawed at 37°C and the cell mixture immediately combined with 1ml warmed media before transferring to the prepared plate. 90% media replacement was carried out the following day to remove dead cells and residual DMSO. Following recovery, cells were monitored daily to assess viability and growth rates.

2.2 Production of his-tagged p16

2.2.1 Transfection

Histidine-tagged vector pcDNA3.1(+) containing a 524bp gene insert for p16 was ordered from Invitrogen GeneArt and amplified in competent bacteria (DH5α) using the GeneJET Plasmid Maxiprep Kit (Thermo Scientific) to produce 1ml purified DNA (589.4ng/µl). For a 6-well culture plate containing 50-70% confluent wild-type cells, DNA vector was added to 100µl appropriate serum-free medium at RT for each well. Transfection agent was added for 15min to allow for complex formation. Cell medium was aspirated from the culture plates and replaced. The transfection complex was added dropwise and the plates maintained as described above for 24h before harvesting. Initial experiments suggested that Lipofectamine resulted in lower transfection efficiency and cell viability than Attractene, and so the latter was used for all further experiments.

SiHa and SAOS-2 (osteosarcoma) wildtype (w/t) cells were transfected with the target vector or a green fluorescent protein (GFP) control insert in pcDNA3.1(+) in order to optimise DNA and reagent levels (Figure 2.1). These cell lines were chosen for their disrupted pRb pathway, thus permitting overexpression of p16 without initiation of cellular senescence and subsequent difficulties in culturing. Reduced levels of Attractene were found to have no adverse effects on transfection efficiency. SiHa cells are capable of producing larger quantities of protein from the same levels of DNA compared to SAOS-2, and so were chosen for further protein purification work. Both p16 expression and controls suggest that optimal reagent and plasmid concentrations can be seen in Table 2.2.

Plate type	**DNA (**ng**)**	**Attractene (**µl**)**	**Media volume (**µl**)**
96 well	0.2	0.5	50
24 well	0.4	1.0	60
6 well	1.2	3.0	100
10cm	4.0	10.0	300

TABLE 2.2: The optimised levels of DNA, Attractene and serum free media used for transfecting His-tagged p16 in pcDNA3.1(+) vector into SiHa wildtype cells.

2.2.2 Purification

Plate capture

Production of his-tagged protein in SiHa cells was confirmed by capture on a nickel-coated 96-well plate (Pierce, Thermo Scientific) with binding capacity of ≈9pM per well. Transfected SiHa cells were lysed with Triton buffer (see Table 2.3) and the lysate (0.8ng/µl or 0.8M) diluted 1:1000 in PBS. 50µl was added to each well required and incubated on a plate shaker at 500rpm for 1h at RT. Wells were washed three times with PBS-Tween and 100µl primary antibody in PBS added at dilutions of 1:2000, 1:1000 and 1:500 in duplicate for 1h at RT. After washing, secondary antibody at 1:1000 in PBS was added for 1h, the plate washed and visualised immediately upon addition of enhanced chemiluminescence (ECL) solution using a photometric plate reader (Fluoroskan Ascent FL, Labsystems). The results can be seen in Figure 2.2.

Bead capture

His-tagged p16 was captured on Ni^{2+} Agarose beads (Ni-NTA Agarose, Qiagen) with 400µl lysate supernatant combined from four confluent 10cm plates of transfected

SiHa. After washing, cells were lysed with NP-40 lysis buffer for 30min on ice and sonicated. The resulting lysate was filter-sterilised with 0.22µm membrane to remove larger molecules or residual cellular debris. Ni^{2+} Agarose beads were washed, pelleted and resuspended three times in Wash Buffer I before suspending in 1ml lysis buffer. The lysate was incubated with the beads for 1h at 4°C on a rotating incubator. The bead-protein complex was run through a prepared 10ml column with 5x Wash Buffer I washes and 3x Wash Buffer II washes before eluting wih 2ml Elution Buffer for 30min at 4°C on a rotating incubator. The elutant was added to a Zeba Spin Desalting Column (7K MWCO, Thermo Scientific) and centrifuged at 1000g for 2min at 4°C to remove salts and other small molecules. The concentration was determined by spectrophotometry of the 260/280nm absorbance ratio to be 33.3µg/ml averaged over three measurements. The purified protein was diluted 4:1 with 80% glycerol in PBS [v/v], aliquoted and snap frozen in liquid nitrogen for storage at -80°C.

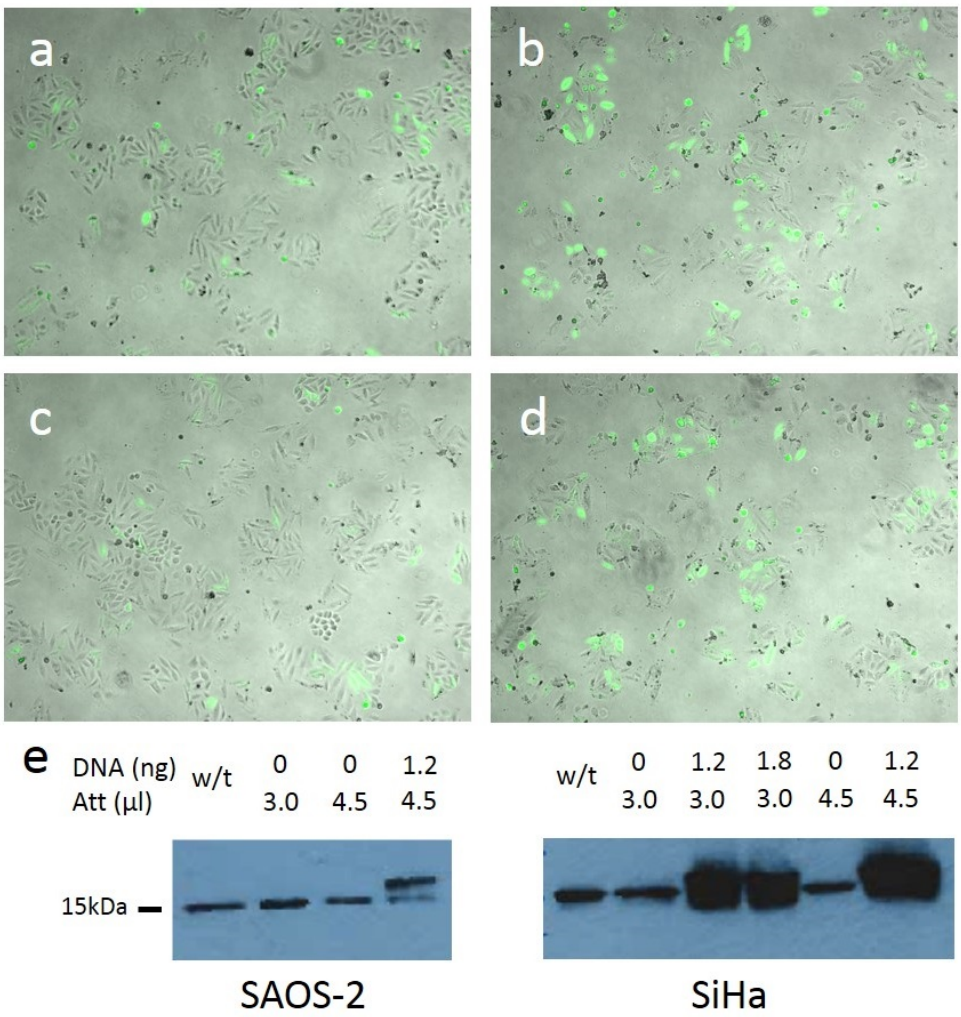


FIGURE 2.1: GFP positive controls in SiHa for a) 1.2 µg DNA, 3.0 µl Attractene, b) 1.8 µg DNA, 3.0 µl Attractene, c) 1.2 µg DNA, 4.5 µl Attractene and d) 1.8 µg DNA, 4.5 µl Attractene. Reduced Attractene does not affect transfection efficiency. e) Demonstration of effective transfection of His-tagged p16 (higher molecular weight) into SAOS-2 and SiHa cells alongside endogenous WT p16 expression in Western blot.

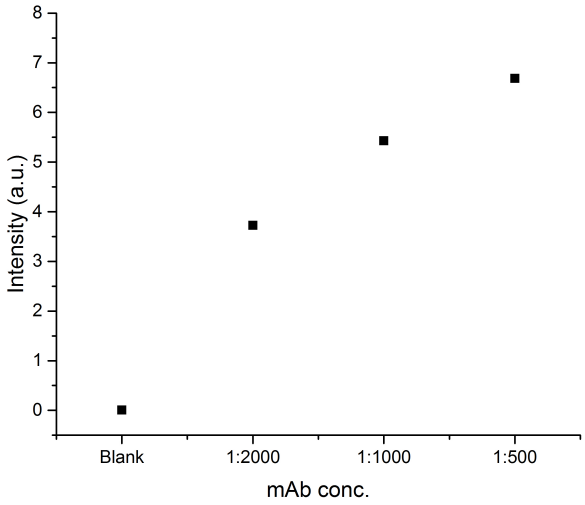


FIGURE 2.2: Lumimetric intensity of nickel plate-captured his-tagged p16 protein from crude cell lysate, demonstrating that the protein can be expressed and extracted using this method. The blank control refers to a block-only plate with 1:1000 primary antibody dilution.

2.3 Western Blotting

For ease of reading, all buffer constituents and gel reagents are listed in Tables 2.3 and 2.4 respectively. All Western Blotting equipment originates from BioRad unless otherwise stated.

Buffer type	Conc.	Constituents
NP-40	1x	0.1% NP-40 [v/v], 25nM HEPES, 150mM NaCl 50mM NaF, 1x PIM, 5mM DTT
Triton	1x	0.10% Triton x100 [v/v], 50mM HEPES (pH7.6) 0.1mM EDTA, 150mM NaCl, 10mM NaF 1x PIM, 2mM DTT
Urea	1x	7M urea, 25mM HEPES, 25mM NaCl 0.05% Triton x-100 [v/v], 5mM DTT
Sample buffer	2.5x	5% SDS [v/v], 25% glycerol [v/v], 0.3M Tris (pH6.8), Bromophenol Blue *Diluted with 1:5 DTT/lysate*
Running buffer	10x	1% SDS [w/v], 30.27g Tris, 144g glycine in 1l dH_2O *Diluted with dH_2O*
Transfer buffer	10x	29g Tris, 144g glycine in 1l dH_2O. *Diluted with 2:7 MetOH/dH_2O*
ECL I	1x	0.1M Tris-HCl (pH8.6), 1ml 250mM Luminol in DMSO, 440µl 90mM p-Coumaric acid in DMSO, in 100ml dH_2O
ECL II	1x	0.1M Tris-HCl (pH8.6), 64µl H_2O_2, in 100ml dH_2O

TABLE 2.3: Constituents of lysis, sample and SDS-PAGE buffers used during the project, particularly for Western Blotting. Lysis buffers were all made up to 2ml with dH_2O. See (1) for details of abbreviations.

2.3.1 Cell lysing and protein quantification

Harvesting

Cells were harvested from culture plates using a cell scraper after three washes with PBS. The resulting 1ml suspension in PBS was centrifuged at 400g for 5min at 4°C and the supernatant removed. Cell pellets were lysed immediately or snap-frozen in liquid nitrogen for storage at -20°C.

Gel type	Constituents
Running gel	5.0ml dH_2O, 6.0ml 30% acrylamide mix
	3.8ml 1.5M Tris (pH8.8), 150µl 10% SDS [w/v]
	150µl 10% APS [w/v], 6µl TEMED
Loading gel	4.1ml dH_2O, 1.0ml 30% acrylamide mix
	750µl 1.0M Tris (pH6.8), 150µl 10% SDS [w/v]
	150µl 10% APS [w/v], 6µl TEMED

TABLE 2.4: Constituents of handcast SDS gels used to separate proteins. Volumes and concentrations described produce 2 12% gels of 1.0mm thickness. See (1) for details of abbreviations.

Lysing

Thawed pellets were resuspended in 100µl freshly prepared lysis buffer and incubated on ice for 15min. The samples were transferred to a microcentrifuge at 16,100g for 15min at 4°C, the protein supernatant transferred to a fresh Eppendorf tube and the cellular debris discarded. Triton, urea and NP-40 lysis buffers were tested with SiHa and CaSki cells (Figure 2.3). Use of urea buffer resulted in an additional band below that associated with native p16, indicating degradation or cleavage. NP-40 lysis buffer was used for all subsequent lysing steps.

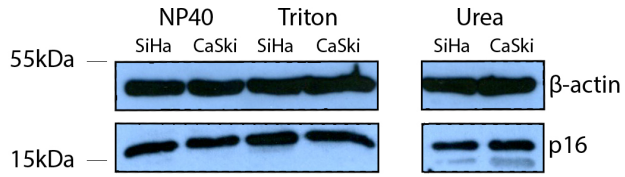


FIGURE 2.3: Expression of p16 and house-keeping gene β-actin in p16-positive cell lines SiHa and CaSki (both HPV16 positive squamous cell carcinoma of the cervix). The use of urea lysis buffer leads to the appearance of lower molecular weight bands.

Bradford Assay

1µl of each lysate sample was added, in duplicate, to 200µl 1x Bradford Reagent (Bio-Rad) in a clear 96-well plate. Protein standards were prepared in the same way using samples of 4.000, 2.000, 1.000, 0.500, 0.250, 0.125, 0.068 and 0mg/ml Bovine Serum Albumin Fraction V (BSA, Melford) in dH_2O [w/v]. Preparations were thoroughly

mixed in a plate shaker at 500rpm for 10min. Colorimetric response was measured using a microplate reader (Wallac Victor^3 1420-050 (Perkin Elmer) or Spark 20M (Tecan)) and the resulting values plotted against the standard BSA curve to give average total protein content per sample in duplicate. Normalised protein volumes were diluted 1:1 with dithiothreitol (DTT) sample buffer and the finalised samples heated at 70°C for 4 minutes to denature the proteins. Samples were stored on ice until loaded.

2.3.2 Gel electrophoresis and imaging

Gel, tank and membrane preparation

12%T (total monomer concentration) SDS gels were produced in gel casts and stored at 4°C for no more than two weeks before use. Upon addition of tetramethylethylenediamine (TEMED, Thermo Scientific) the running gel mixture was pipetted rapidly into the gel casts and left to set. This was repeated for the loading gel, with a well comb of appropriate size and thickness inserted for sample loading. Finished gels were loaded into gel cassettes and placed in a Mini-PROTEAN tank filled with SDS running buffer to the height of the gel cast. Loaded protein samples were separated on the gel at 160V for 1h.

Nitrocellulose membrane and blotting paper were soaked in transfer buffer prior to use. The protein bands were transferred to the nitrocellulose membrane using a Mini Trans-Blot Cell at 100V for 1 hour or at 25mA overnight. An ice block was added to the transfer tank to prevent voltage fluctuations due to overheating.

Antibody incubations and visualisation

Transferred membrane was fixed and stained with 4% paraformaldehyde (PFA) in PBS [v/v] and black ink (Pelican) for 15min. Unbound areas of the membrane were blocked using 5% fatless milk powder (Marvel, 'milk solution') in PBS-Tween [w/v] for 1h at RT. The membrane was partitioned using a scalpel according to the weight of the proteins of interest. The relevant sections were incubated for 1h with primary antibody (to protein of interest or housekeeping gene) diluted in milk solution. The membranes were then washed in triplicate with PBS-Tween to remove unbound antibody. This step was repeated for horseradish peroxidase (HRP)-conjugated secondary antibody (Polyclonal Rabbit Anti-Mouse HRP, DAKO). Chemiluminescence was activated by 1:1 ECLI/ECLII blotting for 1 minute, and the bands visualised on Hyperfilm ECL (Amersham) developed under darkroom conditions with a Konica Medical Film Processor (SRX-101A).

2.4 Immunofluorescence

2.4.1 Cell and slide preparation

Cells were seeded onto ethanol-sterilised glass cover slips at appropriate dilution to reach 50-80% confluency after 24 hours (typically 1:4 in 6-well culture plate) unless transfection was required. After washing in triplicate with PBS, cells were fixed for 15min in 4% PFA in PBS [v/v], washed again with 0.1% PBS-Tween [v/v], permeabilised with 0.1% PBS-Triton X100 [v/v] for 15min and blocked for 30min with 3% BSA in PBS-Tween ('blocking buffer'). Cover slips were transferred to a dark box and incubated for 1h at RT with 1:250 primary antibody to p16 (BD Pharmingen Mouse Anti-Human mAb 0.5mg/ml) in blocking buffer, washed with PBS-Tween and incubated under the same conditions with fluorescent secondary antibody (AlexaFluor 488 Goat Anti-Mouse 2mg/ml). Cover slips were then incubated with 4',6-diamidino-2-phenylindole (DAPI, Thermo Scientific), 1:2500 in PBS for 5min for nucleus visualisation, before washing once with PBS-Tween and twice with dH_2O to remove residual saline. Cover slips were mounted on glass slides using Fluorescence Mounting Medium (DAKO) and stored at 4°C overnight before imaging.

2.4.2 Fluorescence imaging and analysis

Images were recorded using the fluorescein isothiocyanate (FITC) and DAPI channels of a Zeiss upright epifluorescence microscope with micrometre stage. The DAPI channel was used for focus, and the FITC channel of target-positive samples auto-exposed first to determine an exposure time which would not cause image saturation. Tiled images were compiled in MicroManager using Multi Acquisition with 10% image overlap. Individual exposures were stitched together in FIJI (ImageJ) using image position metadata and the Grid/Collection Stitching plugin.

Nuclear expression of p16 protein was assessed over cell populations by mapping a thresholded nuclear mask from the DAPI channel onto the FITC channel and using particle analysis to assess the mean 16-bit greyscale value of each nucleus. The size (μm^2) and circularity of the assessed nuclei were adjusted until all non-overlapping cells were incorporated. Relative p16 expression was visualised using expression histograms, produced in Origin 2016 - a typical example can be seen in Figure 2.4.

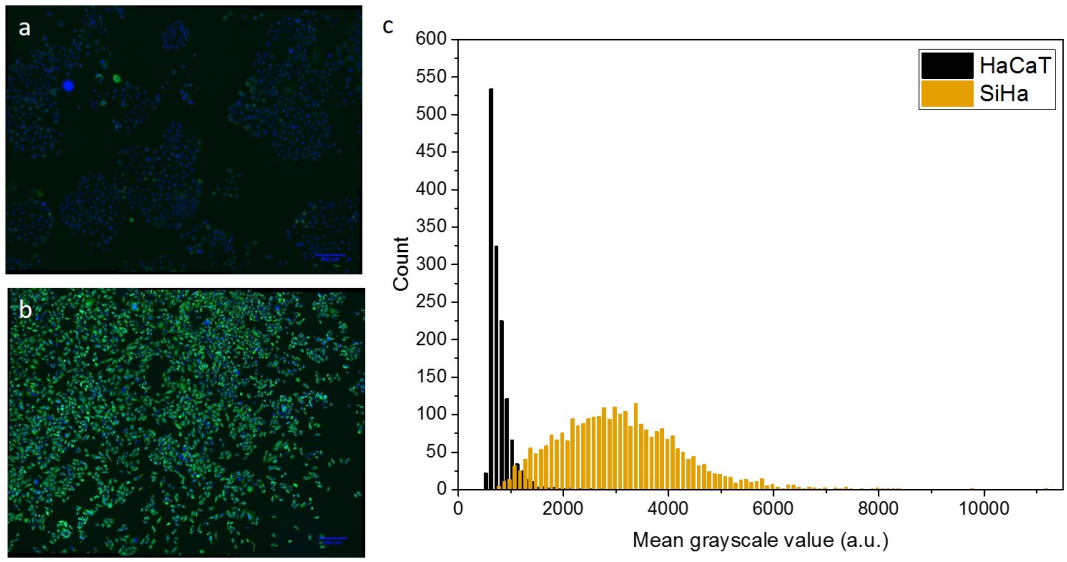


FIGURE 2.4: Tiled immunofluorescence images with p16 expression in green and nuclear staining in blue of a) HaCaT cells, b) SiHa cells. c) A population histogram of the corresponding nuclear expression of p16 in both cell types across the full 16-bit grayscale range.

2.5 Enzyme-linked immunosorbent assay (ELISA)

ELISAs were used to calculate the assay detection sensitivity of antibodies to p16 (BD Pharmingen mAb, mouse anti-human 0.5mg/ml). Purified p16 (p16-INK4a, Recombinant Human, RayBiotech) was reconstituted at 10ng/µl in PBS and the protein solution incubated in 0.1M sodium carbonate buffer such that the total mass per well was 100.00, 50.00, 25.00, 12.50, 6.25, 3.13, 1.56 and 0.78ng in a 50µl volume. Once incubated for 1h at RT on a plate shaker at 500rpm, the wells were washed three times with 0.1% PBS-Tween [v/v] and blocked with 200µl 3% BSA in PBS-Tween [w/v] for 1h at RT. An additional block-only row was used for an average blank value. Blocking buffer was replaced with primary antibody diluted 1:1000 in blocking buffer, incubated for 1h, washed again and the process repeated for HRP-conjugated secondary antibody (Polyclonal Rabbit Anti-Mouse HRP, DAKO). After washing, expression was visualised immediately using 50µl ECL solution on a photometric plate reader. The initial ELISA for BD Pharmingen mAb 0.5mg/ml can be seen in Figure 2.5. This was repeated in triplicate in Chapter 4 Section 4.5.2.

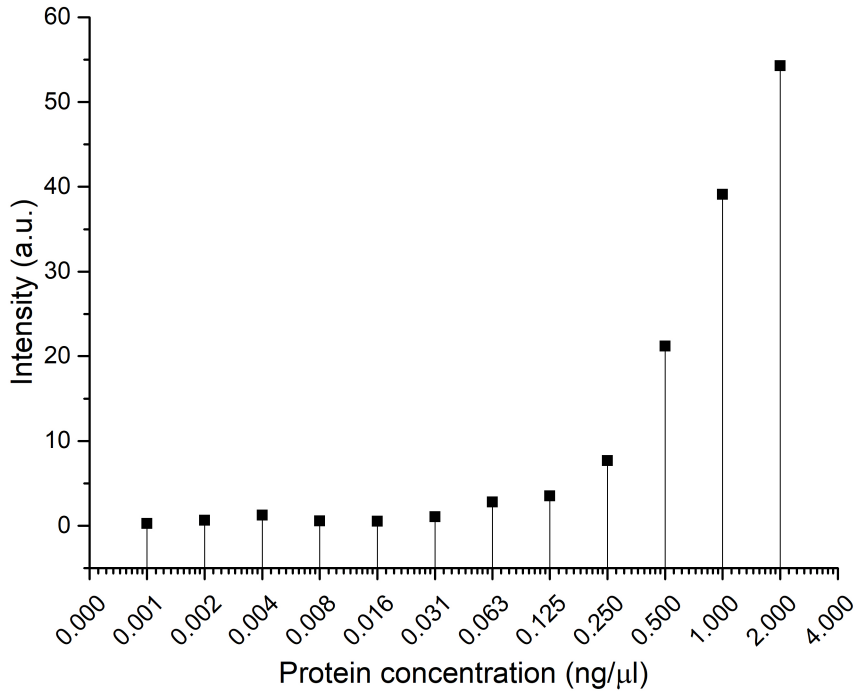


FIGURE 2.5: Lumimetric intensity of BD Pharmingen mAb to p16, diluted 1:1000, when visualised with HRP-conjugated secondary antibody. The ELISA indicates an antibody limit of detection in the region of 0.125ng/µl for this assay format. Antibody detection sensitivity is an important consideration for conjugated nanoparticles.

2.6 Paper-based immunoassays

2.6.1 In-house dipstick assays

Basic in-house dipstick assays were produced by cutting a nitrocellulose membrane (HiFlow Plus, 130µm membrane thickness) into 5mm by 35mm strips (Assay type 'A'). Paper-based devices were also constructed using a nitrocellulose membrane with adhesive backing (Hi-Flo Plus 180 Membrane Card, Merck Millipore) attached to a cellulose wicking pad (C083 Cellulose Absorbent Sample Pad Roll, SureWick, Merck Millipore) at the top and absorbent glass fibre pad (GFDX Glass Fiber Conjugate Pad Roll, SureWick, Merck Millipore) at the base (Assay type 'B'). In both cases, 0.5µl purified protein (p16-INK4a, Recombinant Human, RayBiotech) in PBS was spotted onto the centre of the strips and air-dried for use in testing for antibody binding activity (Figure 2.6).

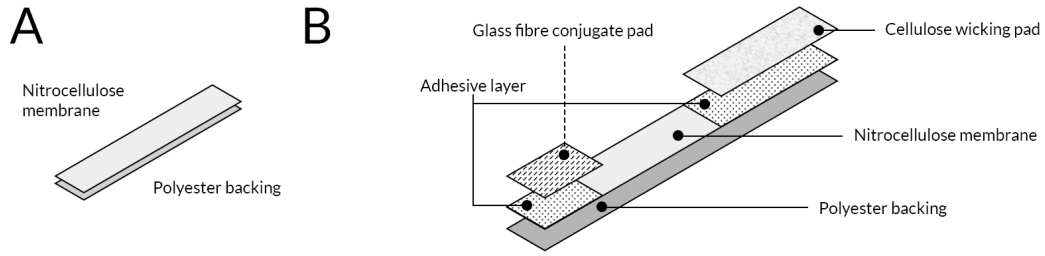


FIGURE 2.6: A) A simple flow device consisting of a backed nitrocellulose membrane. B) An improved architecture which facilitates conjugate uptake and storage through the glass fibre pad, and increased volume and capillary force through the addition of a cellulose wicking pad.

2.6.2 Commercial conjugation assays

Commercial dipstick assays ('Conjugate Check & Go!', Expedeon) were used to test for successful conjugation. In addition to a nitrocellulose sample strip (400µm membrane thickness) and cellulose wicking pad these assays have a printed test line of Protein A/G mix which binds to the fragment crystallizable (Fc) region of most species of monoclonal antibody, including non-reactive isotype controls. Commercial assays were also used as a vehicle for purified protein spotting where stated.

In all cases, the base of the prepared dipstick assay was placed in a clear 96-well plate containing a homogeneous mixture of 20µl nanoparticle sample and 20µl 'conjugate buffer' + 1% BSA [w/v] (Figure 2.7). Conjugate buffer 10x concentrate was supplied with the assay kit, prepared according to the recommended method and stored at 4°C. Paper assays were placed backing down at a slight angle from the vertical to prevent uncontrolled solution wicking caused by surface contact with the well chamber. These were left for 10-20 minutes until the assay was saturated and no liquid remained in the well. Assays were photographed using a dSLR camera (Nikon D3300) or scanned with an Epson 4490 Photo for colorimetric inspection.

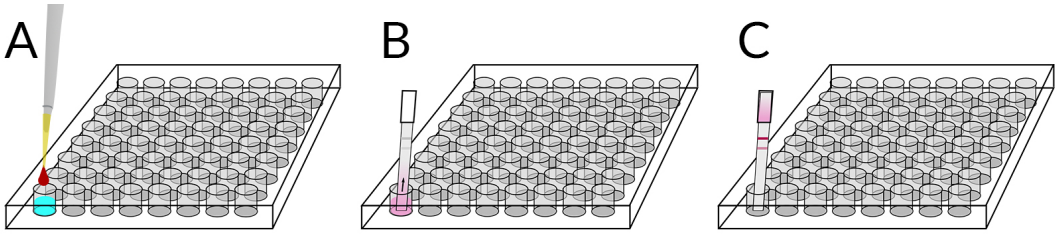


FIGURE 2.7: A pictorial representation of the dipstick method. A) Conjugates are applied to a buffer solution, either in isolation (for testing purposes) or to cell lysate. B) The solution is drawn up an inserted nitrocellulose assay strip (Figure 2.6 by capillary action. C) The conjugates interact with a dried target line (either purified p16 protein or a binding ligand for the same, when using lysate). Excess unbound conjugates dry into the wicking pad.

2.7 Nanoparticle blotting

Nitrocellulose membrane (BioTrace NT) of thickness 101.6µm and pore size 0.2µm was blotted with purified protein (p16-INK4a, Recombinant Human, RayBiotech) in PBS and left to dry. The membrane was then blocked with 5% BSA or fatless milk powder in PBS-Tween for 1h at RT or overnight at 4°C. After rinsing with PBS-Tween, a conjugate solution was applied to the surface for 1-2 hours. In the case of blank nitrocellulose membrane, typically 2ml conjugate would be applied to a 6x8cm surface and left on a rocking plate to prevent solution pooling. For wax-printed 96 well plate designs on membrane, 20µl conjugate was added to each circle (Figure 2.8). Membranes were then soaked in PBS-Tween for 1h to remove non-specifically dried conjugates, before being photographed for colorimetric inspection.

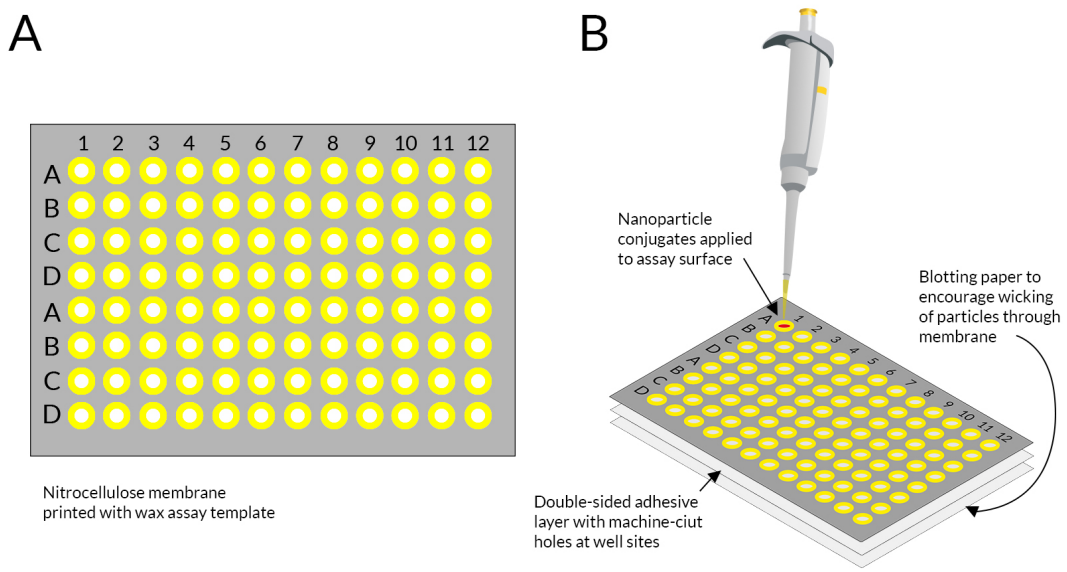


FIGURE 2.8: A) Plan view of the wax-printed '96-well' assay, in which resistive wax is patterned onto a nitrocellulose membrane. B) The target or capture ligand is dried into the exposed 'wells' of the membrane, before the conjugates are applied. Conjugates may either bind to the target in question or flow through to a blotting layer underneath.

2.8 SERS mapping

SERS mapping with conjugate-incubated cells was carried out using a Renishaw inVia confocal Raman microscope equipped with a 500mW 785nm laser and 50x magnification lens. Brightfield images were recorded for hotspot co-registration. Spectra were acquired using a 1200l/mm grating, 10% laser power with edge filter and no pinhole, spectra centred on 1400cm^{-1} and a step size of 2µm unless otherwise stated. Prior to sample measurement, a silicon standard was measured using 100% laser power and 1s single acquisition. The characteristic vibrational peak at 520cm^{-1} was manually calibrated as required.

After acquisition, spectra were smoothed using a 7-point Savitzky-Golay filter and baselined using the proprietary automated baseline tool in Wire 4.0 or 4.2. Normalisation was carried out against the peak intensity of the silicon reference.

Chapter 3

Wavelength modulated Raman spectroscopy for cell classification

3.1 Introduction

One of the major technical barriers to the introduction of Raman spectroscopy as a clinical tool is the rarity of inelastic photon scattering events. The acquisition of high-quality spectra is further compounded by the presence of autofluorescence in samples, which is a particular problem for biological interrogation and can often lead to the loss of smaller Raman peaks behind a fluorescent background. Several approaches for reducing the background have been explored previously, including post-acquisition baseline correction, stimulated Raman spectroscopy (SRS) and time-, frequency- and wavelength-domain methods [127].

Time-domain methods are based on the lifetime differences between emitted fluorescent and Raman scattered photons, on the order of nanoseconds and picoseconds respectively. As such it is a multimodal analogue of fluorescence lifetime imaging (FLIM). The use of an ultrafast laser pulse and temporal gate allows the almost instantaneous Raman scattering event and slower fluorescence relaxation to be separated in time. Temporal gates include Kerr gates [128], intensified CCD or CMOS devices [129, 130] and streak cameras [131]. Such methods, which require high-energy laser sources to generate a detectable signal in short pulses, are typically not suitable for use with biological samples due to the risk of photodamage [129].

Frequency or phase-modulated approaches include frequency demodulation [132] and phase-nulling approaches [133]. In the former, a high-frequency sinusoidal waveform applied to the laser source causes differences in the phase shift and subsequent amplitude of Raman and fluorescence events. This results in a suppression but not necessarily complete removal of the autofluorescent background.

Wavelength modulated Raman scattering (WMRS) is a wavelength-domain approach, along with shifted excitation Raman difference spectroscopy (SERDS) [134] and subtracted shifted Raman spectroscopy (SSRS) [135]. WMRS has been used to investigate single living Chinese Hamster Ovary cells (CHO-K1) [136] as well as T lymphocytes ($CD4^+$, $CD8^+$), killer T cells ($CD56^+$) and dendritic cells ($CD303^+$, $CD1c^+$) [137]. It has also been shown to provide high specificity and sensitivity for the characterisation of human urothelial (SV-HUC-1) and bladder cancer (MGH) cell lines cultured normally [138] or exposed to urine [139], which represents another step towards an automated optical platform for cancer screening. However, molecular markers of pathologies vary from disease to disease, necessitating for this thesis an in-depth study of cells specific to the cervical cancer context.

3.2 Concept: wavelength modulation

During a WMRS measurement, the laser wavelength is modulated throughout the acquisition process using a tunable laser [138]. This causes the Raman spectral peaks to shift fractionally when measured in terms of scattered wavelength. When this shift is measured across several wavelengths, a 'differential spectrum' can be generated which demonstrates the true Raman peaks in wavenumber shift from the incident wavelength. This wavelength modulation does not affect the position of autofluorescence from the sample and substrate, allowing the background to be automatically removed.

Automatic autofluorescence removal is conducted on the basis that fluoresence background is static in reference to a shift in incident wavelength. An individual spectrum S_j measured at a wavelength shift $\Delta\lambda_i$ of the tunable laser can be represented by the superposition of the Raman S_R and fluorescent S_F parts, where $S_j(\lambda_i) = S_F(\lambda_i) + M_{ik}(\Delta\lambda_i)S_R(\lambda_k)$. M_{ik} is the fractional shift matrix, determined by the modulation of wavelength. The separation of the Raman and fluorescence components is achieved by representing M_{ik} using a lock-in procedure [136], where the modulated Raman spectrum can be elucidated from a 'noisy' fluorescence spectrum. Numerous methods have been tested with wavelength modulation for the removal of this static fluorescence background, including standard excitation Raman difference spectroscopy, standard deviation, Fourier filtering, least-squares fit and principal component analysis (PCA), with PCA providing the best signal-to-noise ratio (SNR) [140]. Using Principal Component Analysis (see Section 3.3.3), the differential Raman spectrum is formed from the first principal component of the modulated data set, corresponding to the component which captures the greatest variance in the data.

In this study, the principle of background-free global Raman interrogation and multivariate analysis was applied to the automatic classification of cervical smear analogues, fixed in PreservCyt to reflect clinical practice. The primary aim was to test the classification capability of the technique, with additional consideration for the practicalities of sample acquisition time and setup.

3.3 Materials and methods

3.3.1 Cell sample preparation

A total of five established cell lines from squamous epithelial tissue were used for this project, reflecting the intention to investigate WMRS as a method of classifying cervical smear samples. CaSki, SiHa, C33A (all squamous cell carcinoma, adult human cervix), HaCaT (immortalised squamous epithelium, adult human skin) and NHEK (normal squamous epithelium, neonatal human skin) were cultured to 80-90% confluency. Culture methods can be found in Chapter 2, Sections 2.1.1 and 2.1.2. These were harvested by trypsinisation, washed in triplicate by centrifugation and resuspension in sterilised phosphate buffered saline (PBS), and fixed in the clinical methanol-based fixative PreservCyt Solution ('PreservCyt', Cytyc). Samples were stored in PreservCyt at -20°C until use. All centrifugation steps were executed at 400g for 5 minutes at 4°C.

Prior to interrogation, fixed cell samples were washed in triplicate by centrifugation and resuspension in sterilised PBS in order to remove traces of fixing solution, as methanol produces strong Raman signatures. 20µl of PBS cell suspension was then transferred to a quartz cavity slide (Figure 3.1), carefully sealed with a 150µm quartz coverslip and inverted. The quartz assemblies were rested for at least 20 minutes prior to interrogation in order to promote cell adherence to the quartz coverslip, which was placed face down on the inverted microscope assembly for measurement.

3.3.2 Spectral acquisition

A confocal Raman spectrometer built and modified for acquisition of both standard and modulated spectra (Figure 3.2) was used to evaluate the fixed cell samples. Raman spectra were generated from individual cells inside the quartz cavity assembly using a 'SolsTis' tunable continuous wave Ti-Sapphire laser (M Squared, maximum power 6W), delivering 150mW power to a 1µm spot in the sample plane through a 40x oil immersion objective (Nikon, NA1.3). The signal was resolved with a Shamrock SR-303i spectrometer (400 lines/mm grating, Andor Technology) and collected with a CCD camera (Newton, Andor Technology) thermoelectrically cooled to -70°C.

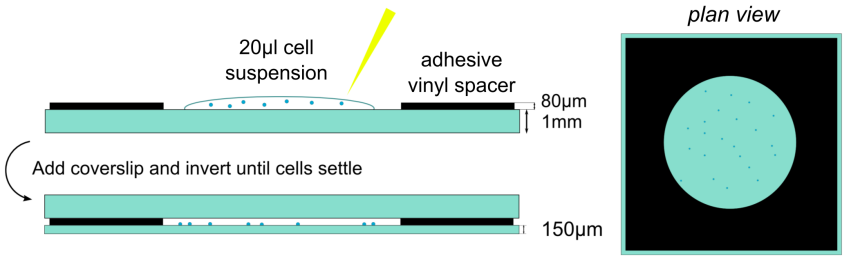


FIGURE 3.1: The cavity slide assembly used for all acquisitions. Slides were sterilised with ethanol, air-dried and polished in between measurements.

'Standard' (spontaneous) and wavelength modulated Raman spectra were acquired using Solis (s) Software (Andor Technology) along with a brightfield microscope image (Imaging Source USB Camera) of each cell.

Standard acquisitions were conducted at 785nm, using 5 accumulations of 2s laser exposure for both the cell and adjacent background. In the modulated acquisitions, 5 accumulations of 2s exposures were performed with five different cycles of excitation wavelengths over a total range 785$\pm$0.5nm. No background acquisition was necessary in the modulated case due to the inherent autofluorescence removal. This approach was justified on two premises. First, the use of five distinct modulated cycles as a form of background removal does not improve the SNR of the raw modulated data to provide an unfair advantage over the standard spectra, which have had their background removed in a different way. Second, the need to acquire two spectra in the standard case (for cell and adjacent background) resulted in a practical acquisition time equivalent to the total modulated acquisition time (50s per data point).

It has been shown previously that no significant phototoxicity was experienced by the cells as a result of laser exposure at similar settings[137].

Reference spectra for laser calibration and spectral normalisation (Section 3.3.3) were acquired from 1µm polystyrene beads in distilled water. These spectra, recorded with a time-stamp either side of each dataset, were used to account for fluctuations in laser power, wavelength and spot location due to heating, fibre-optic coupling or mirror alignment.

3.3.3 Spectral processing and data analysis

All spectra were considered in the fingerprint region of interest, 800 to 1800cm^{-1}. Normalisation was conducted by linearly interpolating the intensity and Raman shift of the characteristic 1001.4cm^{-1} polystyrene peak according to the sample measurement

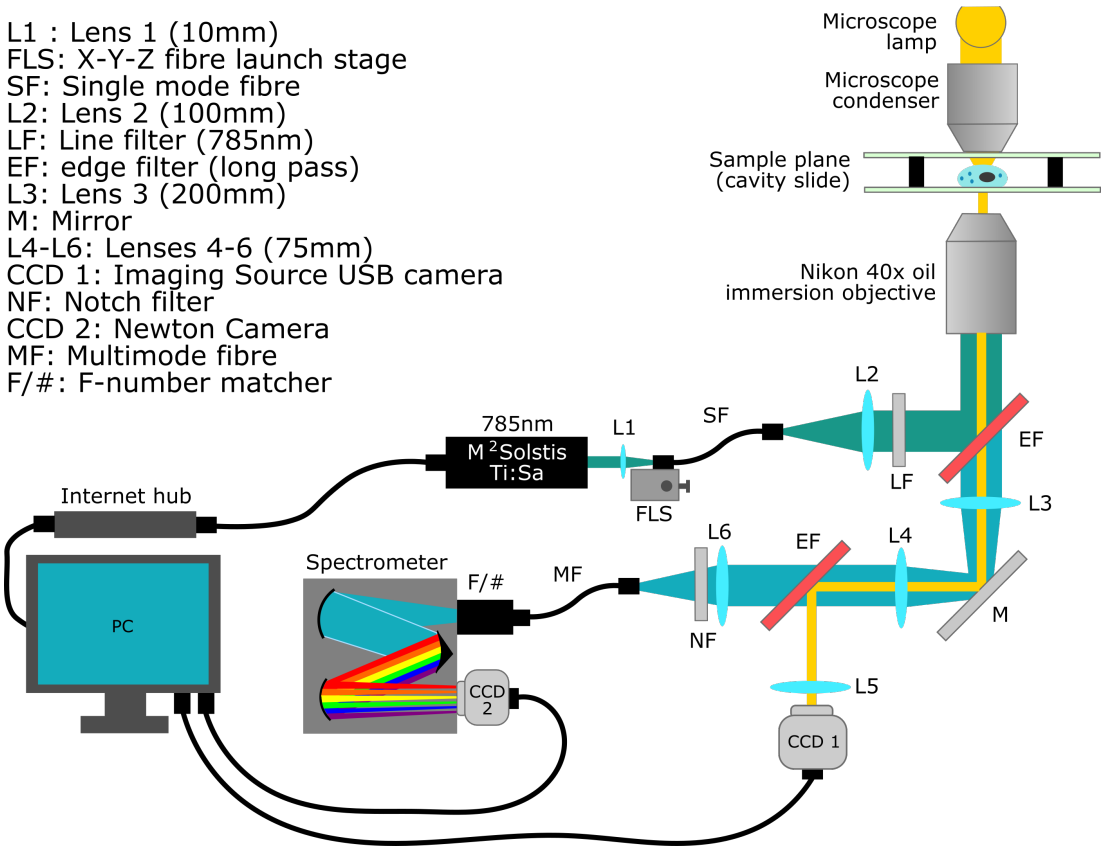


FIGURE 3.2: The optical diagram of the spectrometer designed for WMRS and standard Raman spectral acquisition. Brightfield microscope images are acquired separately on CCD1 and blocked from the spectrometer by secondary EF.

time stamp. This accounts for laser power fluctuations over time, thereby preventing the PCA script from grouping cells according to laser power or wavelength drift. Immediately prior to analysis, individual spectra were assessed and removed from the analysis if a high spectral background was apparent. Such artefacts arise from suboptimal focusing or cell movement during acquisition.

All spectra were analysed using Principal Component Analysis (PCA), a commonly used multivariate spectral analysis technique. PCA breaks spectra down into orthogonal spectral components ('loadings' or 'principal components', PCs) which describe the variation in the dataset. The first principal component describes the greatest variation, with higher components eventually describing noise [141]. Seven PCs were used in all data analysis conducted but a maximum of three displayed for ease of visualisation. Previous research with human immune cells with the same setup indicated that the first seven PCs account for the major variation in the dataset [137], and so this was used for all test models. The weighting or 'score' of these PCs (the extent to which each PC represents a particular spectrum in the dataset) can be used to generate PCA cluster plots, which may be used for diagnostic classification. The loadings and scores represent, respectively, the eigenvectors and eigenvalues of the data covariance matrix, consisting of samples and their spectra at each wavenumber.

3.3.4 Cross-validation of spectral classification

Sensitivity and specificity values can be generated through leave-one-out cross validation (LOOCV) also known as a jacknife [142], described visually in Figure 3.3. Here, a single spectrum is retained and the remaining spectra used to form a training set. The retained spectrum is then classified according to the first seven principal component loadings of the training set using a k-nearest-neighbours (KNN) algorithm [137] with a known number of clusters, equal to the number of cell types used. In a clinical setting, this would require large training sets with clusters assigned to designated test outcomes, for example normal, mild dyskaryosis, and so on.

This is repeated for each individual spectrum, with the collated data displayed as a confusion matrix. In confusion matrices, correctly classified types appear on the leading diagonal, with mis-classifications appearing in off-diagonal positions. Matrices are square with dimensions equal to the number of subsets. The off-diagonal populations indicate which subsets are more likely to be mistakenly classified as another set, potentially allowing for conclusions to be drawn regarding biological similarity or disparity.

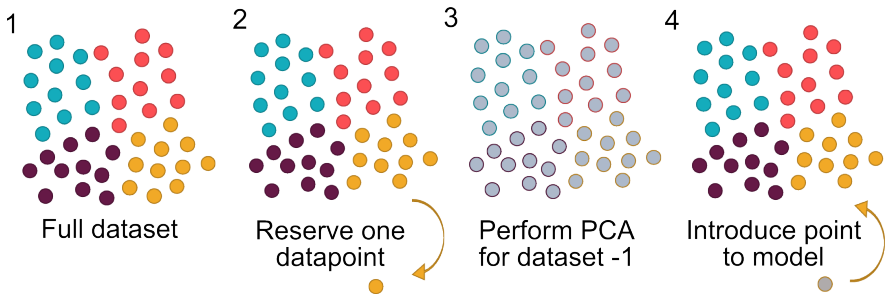


FIGURE 3.3: Pictorial representation of the leave-one-out cross validation method, used to assess the accuracy of the PCA model for classification.

When performed for two datasets, the classifications can be quantitatively estimated as sensitivity and specificity values by assigning the cell line as 'positive' or 'negative'. Equations 3.1 and 3.2 determine the results.

	Sample 1 (Predicted)	**Sample 2 (Predicted)**
Sample 1 (Actual)	True Negative (TN)	False Positive (FP)
Sample 2 (Actual)	False Negative (FN)	True Positive (TP)

TABLE 3.1: The classification of assigned positive and negative samples used to generate sensitivity and specificity data, based on actual and PCA-classified sample type.

$$\text{Sensitivity} = \frac{\text{Test true positive samples}}{\text{All truly positive samples}} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3.1}$$

$$\text{Specificity} = \frac{\text{Test true negative samples}}{\text{All truly negative samples}} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{3.2}$$

Statistically significant variations between spectral intensities were also estimated for each cell line pair using student's t-test at a significance level of $p < 10^{-12}$. The full processing and analysis script used was written in MATLAB by the Dholakia group (University of St Andrews).

3.4 Comparison of human papillomavirus (HPV) status

A preliminary comparison of standard and wavelength modulated Raman spectroscopy as a classification tool was made using fixed C33A (HPV negative) and SiHa (HPV16 positive) cervical carcinoma cell lines, kindly donated by the Scottish HPV Archive [143]. The samples were resuspended in PBS and prepared for analysis as described

in Section 3.3.1. Initial tests with the C33A cell line found that an acquisition time of 2s over five accumulations yielded superior standard Raman spectra to a single acquisition of 10s. The brightfield cell images were morphologically similar, consisting of $\sim$10µm diameter cells which were highly rounded due to the trypsinisation process.

The PCA plots for the standard and modulated Raman can be seen in Figure 3.4 and 3.5 respectively. Spectra can be seen in Figure 3.6. LOOCV was applied to the datasets to yield pairwise sensitivity values of 88.0% and 100.0% for standard and modulated Raman respectively, with specificities of 90.5% and 95.8% respectively. It is clear from both PCA and classification accuracy that modulated Raman spectroscopy provides an improved separation and classification of two cervical carcinoma cell lines with differing HPV status. The main spectral disparities occur at approximately 730, 760, 890, 920, 1130, 1250, 1380 and 1480cm^{-1}. In the scores plot, the PC1/2 projection provides the most distinct separation of samples, as expected.

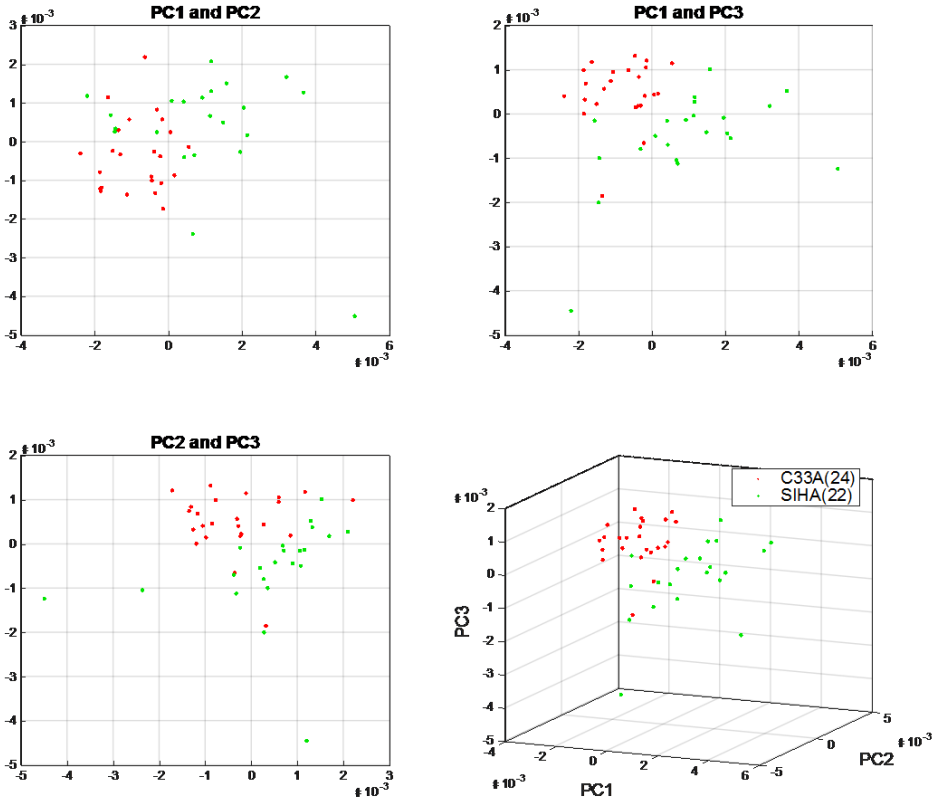


FIGURE 3.4: Standard spectral datapoints represented as PC weightings for C33A (HPV negative carcinoma) and SiHa (HPV-16 carcinoma) cells. 24 C33A and 22 SiHa spectra were used for the comparison.

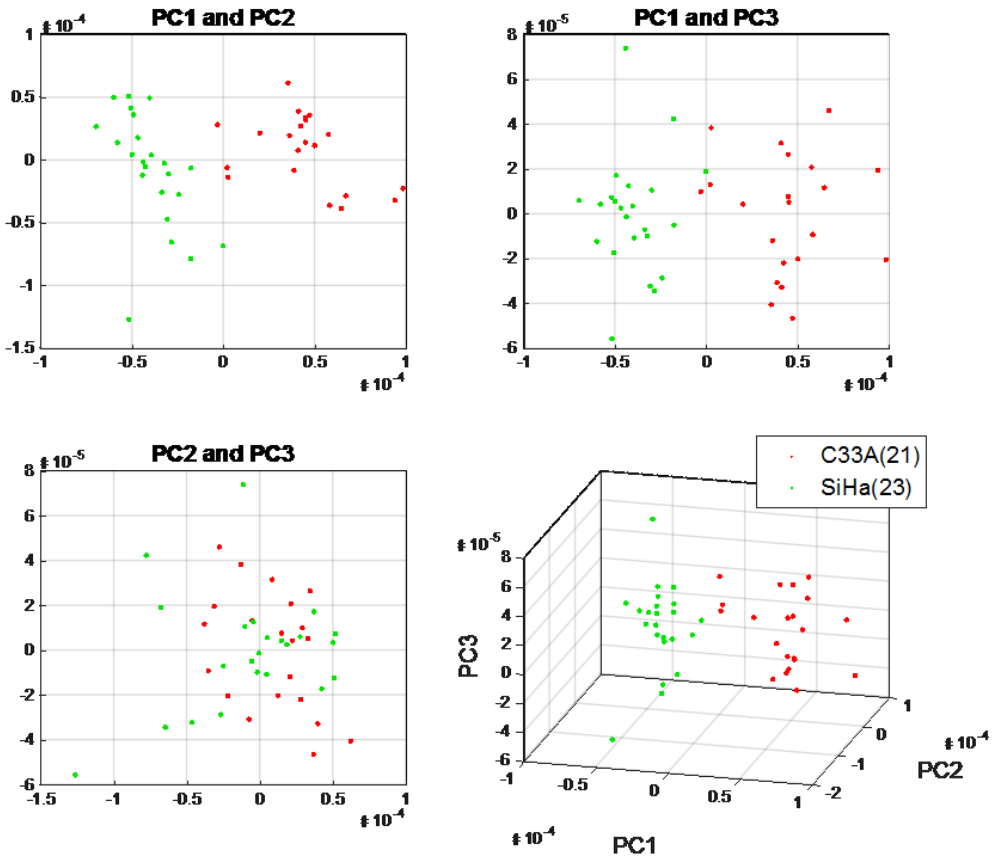


FIGURE 3.5: Modulated spectral datapoints represented as PC weightings for C33A (HPV negative carcinoma) and SiHa (HPV-16 carcinoma) cells. 21 C33A and 23 SiHa spectra were used for the comparison.

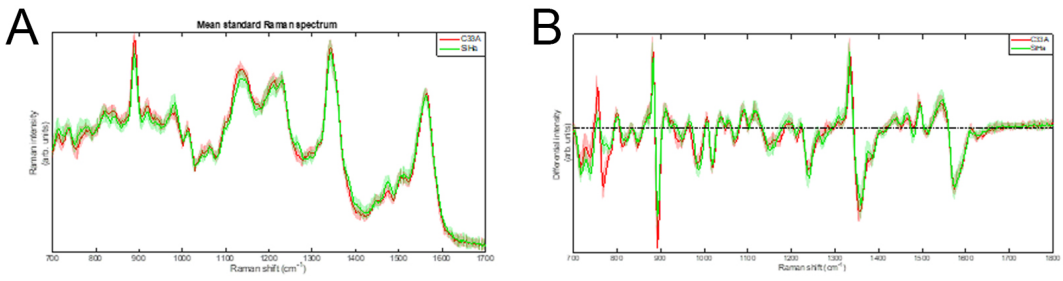


FIGURE 3.6: Average A) standard and B) modulated spectra for C33A and SiHa cells.

3.5 Comparison of transformation status

3.5.1 Comparison of standard and modulated spectral classification by PCA

The investigation of standard and modulated Raman spectroscopy was then extended to four cell lines (one normal, one immortalised, two cancerous). A total of 163 cells were interrogated using both standard and wavelength modulated Raman spectroscopy, consisting of 42 NHEK, 40 HaCaT, 41 SiHa and 40 CaSki cells. As before, the established cell lines in particular could not be discerned on a purely morphological basis, which furthers the case for biochemical classification as a complementary technique. It must be noted that the morphology of the primary NHEK cells varied significantly in terms of both shape and size compared to the established cell lines. Following removal of suboptimal spectra according to Section 3.3.3, Raman spectra and PCA plots were generated as seen in Figures 3.7 and 3.8 respectively. The ability of PCA to act as a classification algorithm was explored using cross-validation of the data. LOOCV was performed on all four cell lines for both datasets; the confusion matrices for standard and modulated PCA can be seen in Tables 3.2 and 3.3 respectively, along with misclassification statistics. For all three established cell lines, the modulated spectra provide better discrimination. However, classification is slightly superior using standard Raman spectra for NHEKs.

3.5.2 Pairwise analysis of cell lines

In order to gain further understanding of both the efficacy of the test and the major vibrational signals contributing to the analysis, each possible pair combination of cell lines were analysed using the PCA script. The sensitivity and specificity values for each of these tests are summarised in Figures 3.9 and 3.10 respectively. The average sensitivity was $84.8 \pm 9.0\%$ and $93.8 \pm 5.2\%$, and specificity was $85.1 \pm 10.7\%$

	Predicted				
Actual	NHEK	SiHa	CaSki	HaCaT	Class acc.
NHEK	**31**	2	1	1	88.6
SiHa	2	**25**	6	7	62.5
CaSki	1	9	**19**	10	48.7
HaCaT	3	9	6	**22**	55.0

TABLE 3.2: The confusion matrix resulting from cross-validation of standard Raman PCA, with accompanying percentage classification accuracy.

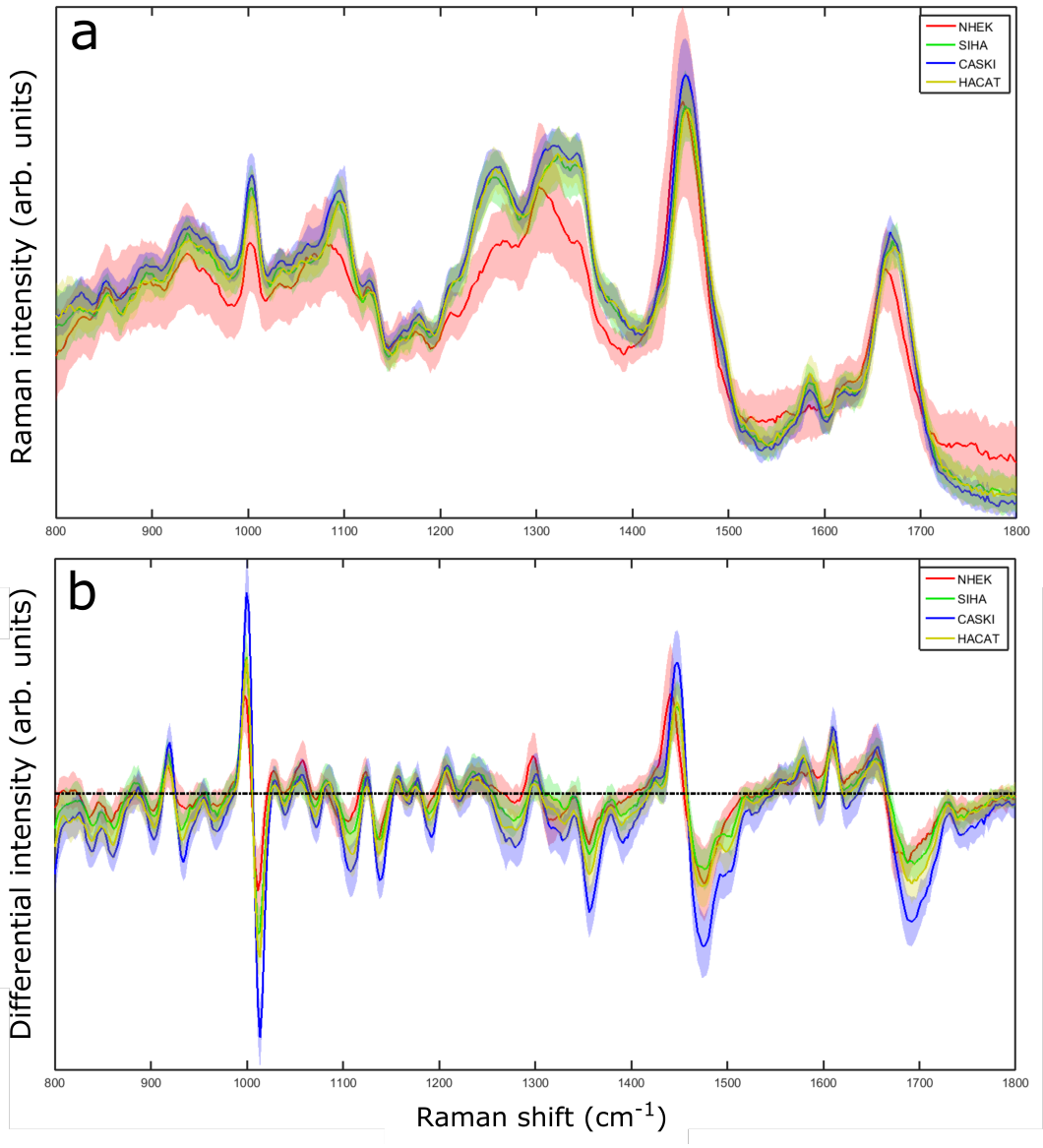


FIGURE 3.7: (a) Standard and (b) modulated Raman spectral averages for the four cell lines used, 800-1800cm^{-1}

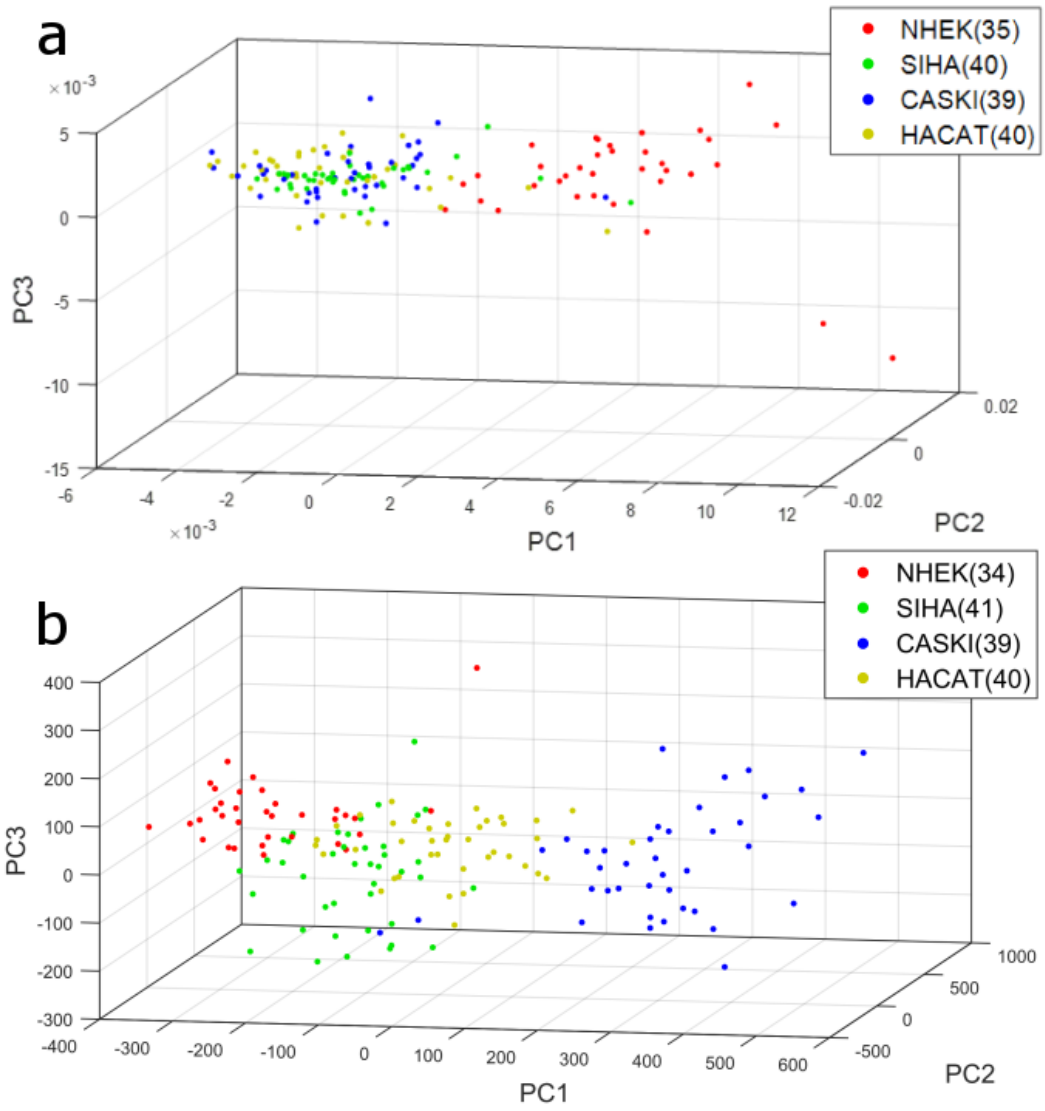


FIGURE 3.8: (a) Standard and (b) modulated PCA weightings for the four cell lines used.

	Predicted				
Actual	NHEK	SiHa	CaSki	HaCaT	Class acc.
NHEK	**29**	4	0	1	85.3
SiHa	1	**30**	0	10	73.2
CaSki	1	1	**38**	0	95.0
HaCaT	0	2	0	**38**	95.0

TABLE 3.3: The confusion matrix resulting from cross-validation of modulated Raman PCA, with accompanying percentage classification accuracy.

and $93.8 \pm 4.3\%$, for standard and modulated PCA respectively, all given as standard error of the mean (SEM). This demonstrates that, for both metrics, wavelength modulated Raman spectroscopy provides superior discrimination as well as lower variation between cell types.

The generated spectra were also marked with significance bands as described in Section 3.3.4. Banded spectra for NHEK-SiHa, NHEK-CaSki and NHEK-HaCaT are shown in Figures 3.11 and 3.12: the remaining combinations (SiHa-CaSki, HaCaT-SiHa, HaCaT-CaSki) did not possess any bands under the t-test statistics used and are omitted. The lack of statistically significant spectral differences for the latter pairings explains the reduced sensitivity and specificity values for these three comparisons. Prior to consideration of significance bands, it can be seen in the standard Raman spectra that for each comparison of NHEKs to the three established cell lines, the same peaks are demonstrating significant change. The peak at approximately 1590cm^{-1} (adenine/guanine) only appears for the immortalised and carcinoma cell lines. The largest peak at approximately 1450cm^{-1} (CH_2 deformation in lipids) is unusual in that it is higher for NHEKs than SiHa or HaCaT cells. All remaining peaks show increases for the cancerous and immortalised cells. The NHEK-CaSki comparison shows a small increase in the C-C skeletal stretching in protein not seen in the remaining comparisons whereas, in the NHEK-SiHa and NHEK-HaCaT studies, the latter cell lines show more pronounced increases in the ring breathing in tyrosine/C-C stretching in proline.

The statistically significant bands and tentative peak assignments based on previous literature [117] are discussed in Section 3.8. Overall, it is clear that the use of a differential spectrum in modulated Raman spectroscopy highlights different statistically significant regions of the spectrum in comparison to standard Raman spectra. This is because the differential method highlights the shape and relative position of the peak, as opposed to just the intensity.

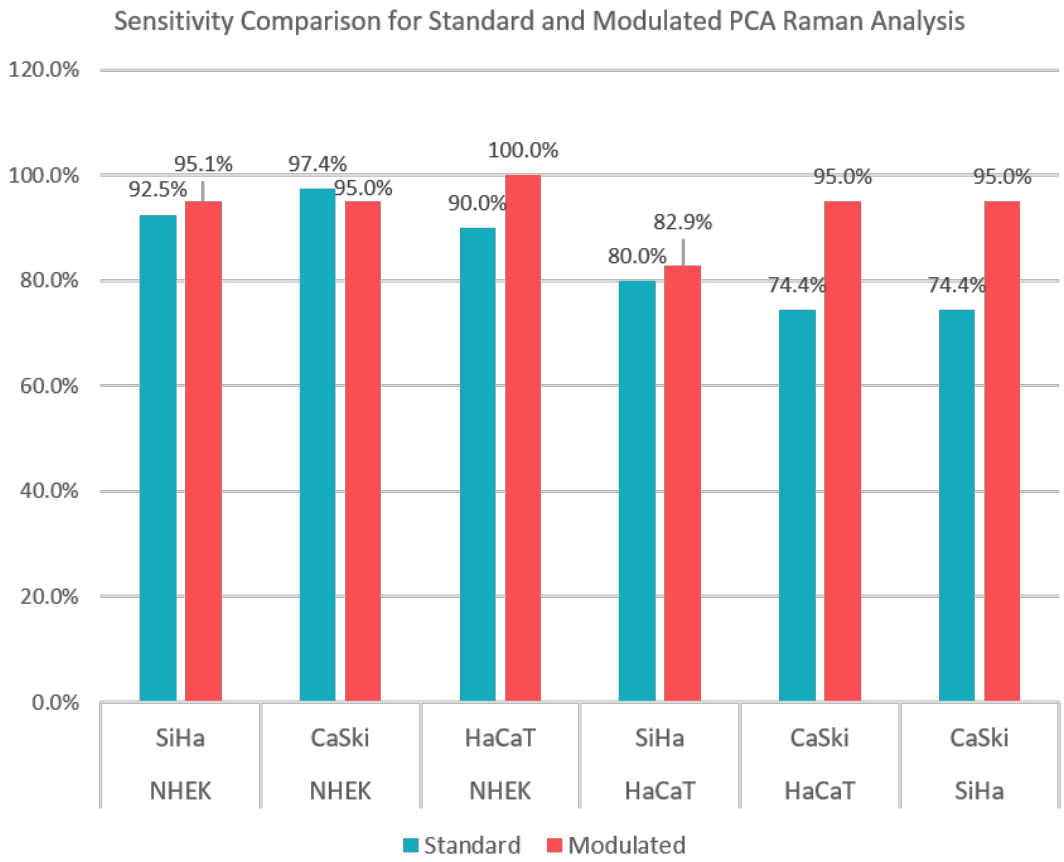


FIGURE 3.9: The sensitivity values for each set of cell lines when cross-validated as a pair.

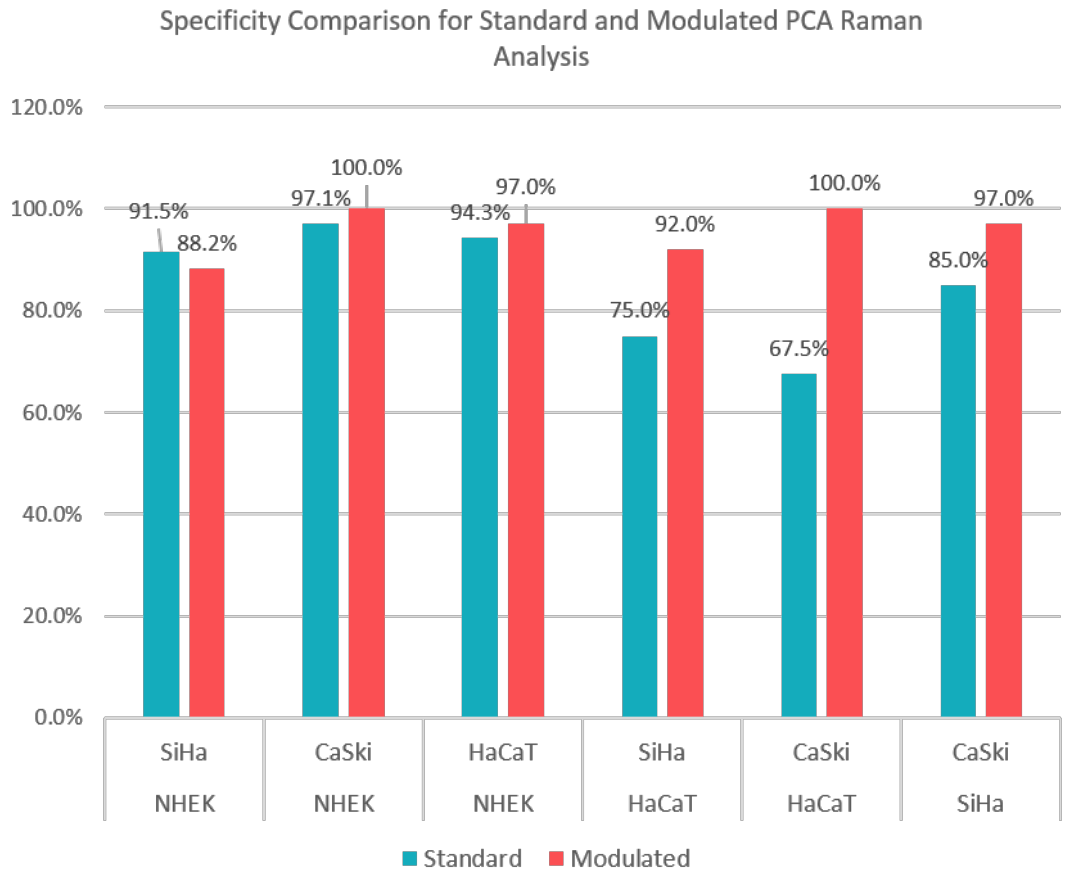


FIGURE 3.10: The specificity values for each set of cell lines when cross-validated as a pair.

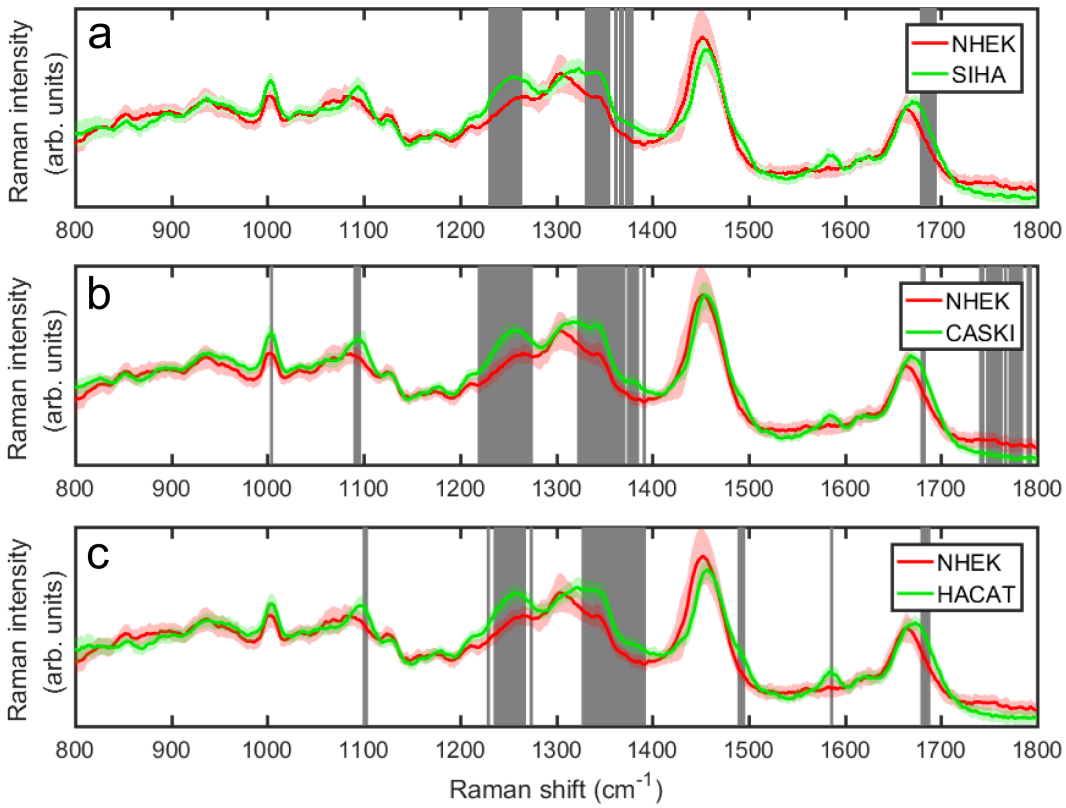


FIGURE 3.11: Standard Raman spectra for the pairwise comparisons demonstrating significant spectral deviations (grey bands) under the t-test statistics used.

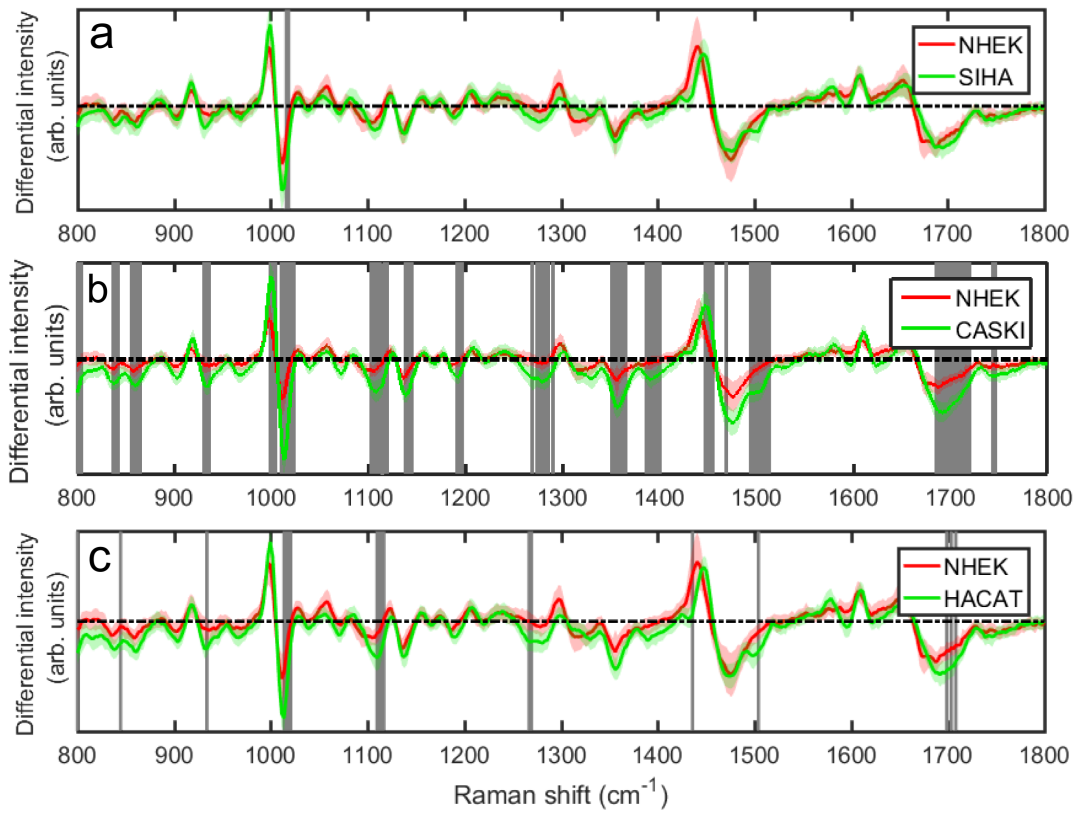


FIGURE 3.12: Standard Raman spectra for the pairwise comparisons demonstrating significant spectral deviations (grey bands) under the t-test statistics used.

3.6 Considerations of sampling location

As discussed in Section 3.1, it was previously found that the Raman spectra for intracellular locations in CHO cells cultured on glass slides were very distinct, with the nucleus providing the optimum SNR. However, this data was generated from adherent cells grown directly onto glass slides, giving the characteristic flattened morphology associated with cells in culture (Figure 3.13). In the case of spectral acquisitions for Sections 3.4 and 3.5, it was not always clear from a combination of camera resolution and morphology whether the nucleus was being targeted.

To investigate the importance of intracellular location on characterisation, further samples of the Section 3.5 cell lines were fixed and each cell interrogated multiple times at locations designated as cell membrane (labelled wall for brevity), cytoplasm, nucleus and nucleolus. Cells were rejected if the specified components could not be discerned clearly. The nucleolus was identified as a high-density sphere within the nuclear volume. All four datasets were recorded in the modulated setup only and analysed as discussed previously.

Figure 3.14 shows the overall PCA plot from the cells, disregarding laser spot location. The resulting confusion matrix classification accuracies are detailed in Table 3.4. Despite some misclassifications (particularly for NHEKs), the average predictive accuracy is high at 91.2%. This confusion matrix can then be expanded to consider location for each cell line in turn. This can be seen for the SiHa line in Figure 3.15, with remaining cell lines in Figures 3.16, 3.17 and 3.18. The classification accuracies for the location-specfic subset of each expanded matrix are summarised in Table 3.5. These are much lower than both classifications by cell line and location-based results from de Luca et al.'s experiment, suggesting that there is insufficient spectral disparity

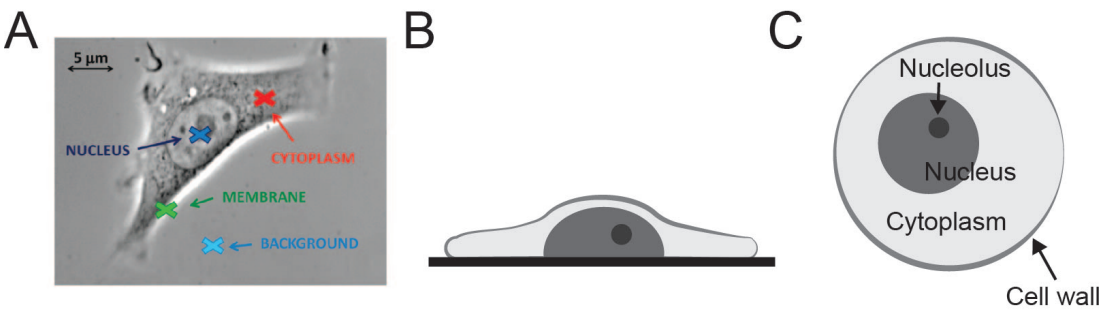


FIGURE 3.13: Intracellular locations of A) the original experiment with CHO cells which indicated classification could be made by the locations shown; B) typical morphology of adherent cells in culture; C) typical morphology of fixed cells following trypsinisation.

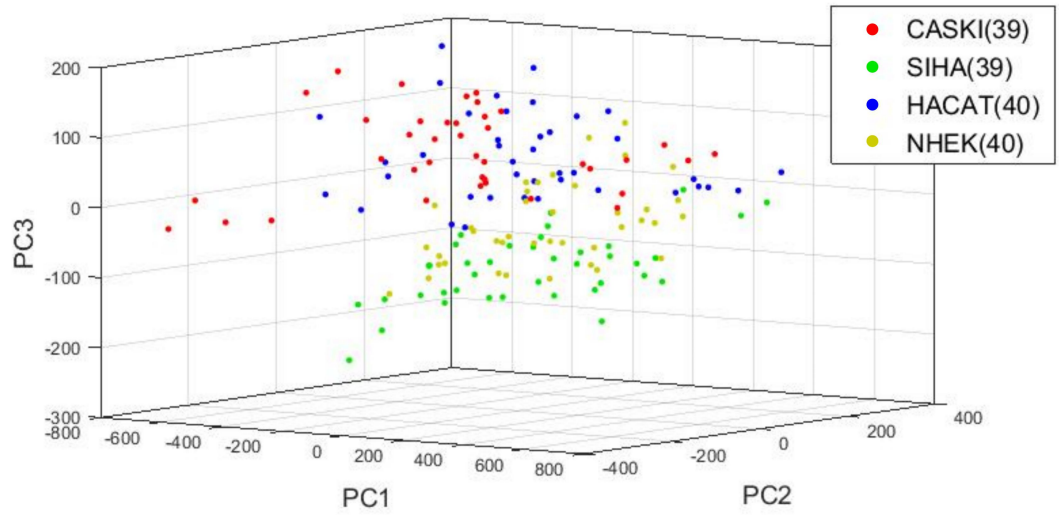


FIGURE 3.14: PCA weightings for the four cell lines classified without consideration of sampling location.

to separate clusters by location.

	Predicted				
Actual	CaSki	SiHa	HaCaT	NHEK	Class acc. (%)
CaSki	**38**	0	0	1	97.4
SiHa	0	**36**	2	1	92.3
HaCaT	1	0	**37**	2	92.5
NHEK	0	2	5	**33**	82.5

TABLE 3.4: confusion matrix resulting from cross-validation of the intracellular location dataset before consideration of internal location. The average classification accuracy is 91.2%.

As such, the matrices were then considered by grouping points according to their location inside (nucleus, nucleolus) or outside (cytoplasm, cell membrane) the nucleus. The summarised classification accuracies shown in Table 3.6. The resulting accuracies from this test indicate that there is sufficient spectral difference to separate these location clusters from inside and outside the nucleus with an accuracy approaching that of cell line classification. However owing to their rounded morphology, the fixed cells do not provide sufficient spectral difference to distinguish between specific cellular regions with the accuracy seen for cells in culture.

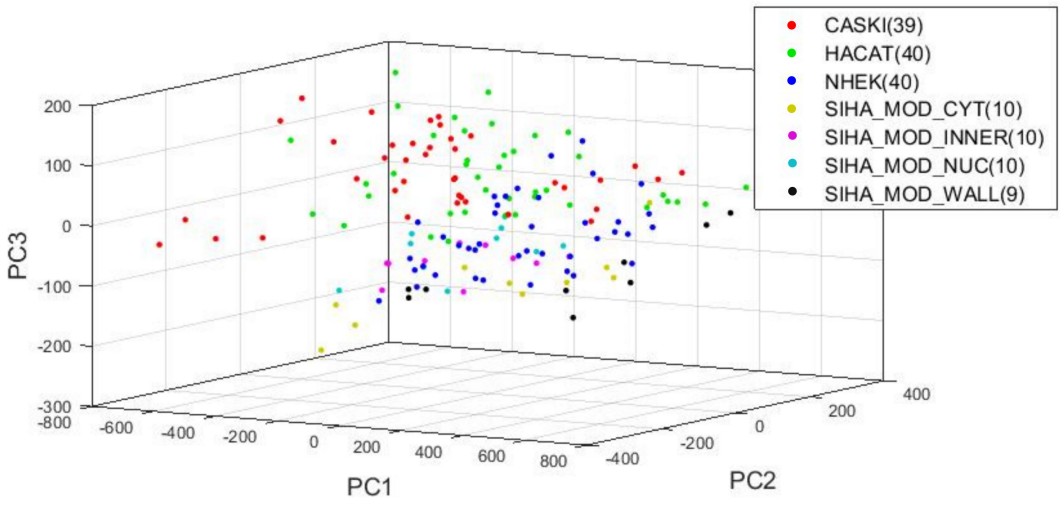


FIGURE 3.15: PCA weightings for the four cell lines classified with consideration of sampling location, subdivided into nucleus, nucleolus ('inner'), cytoplasm ('cyt') and cell membrane ('wall').

	Classification accuracy				
Location	NHEK	SiHa	CaSki	HaCaT	Average
Nucleus	50	60	50	60	55±5
Nucleolus	50	20	50	80	50±21
Cytoplasm	44	88	44	80	64±20
Cell wall	60	50	60	80	63±11

TABLE 3.5: A summary of the classification accuracies resulting from cross-validation of modulated Raman PCA of intracellular regions.

	Classification accuracy				
Location	NHEK	SiHa	CaSki	HaCaT	Average
Inside	93.3	90.0	90.0	94.1	91.9±1.9
Outside	88.9	86.7	77.8	100.0	88.4±7.9

TABLE 3.6: Classification accuracies resulting from cross-validation of intracellular sampling with reduced spatial sensitivity.

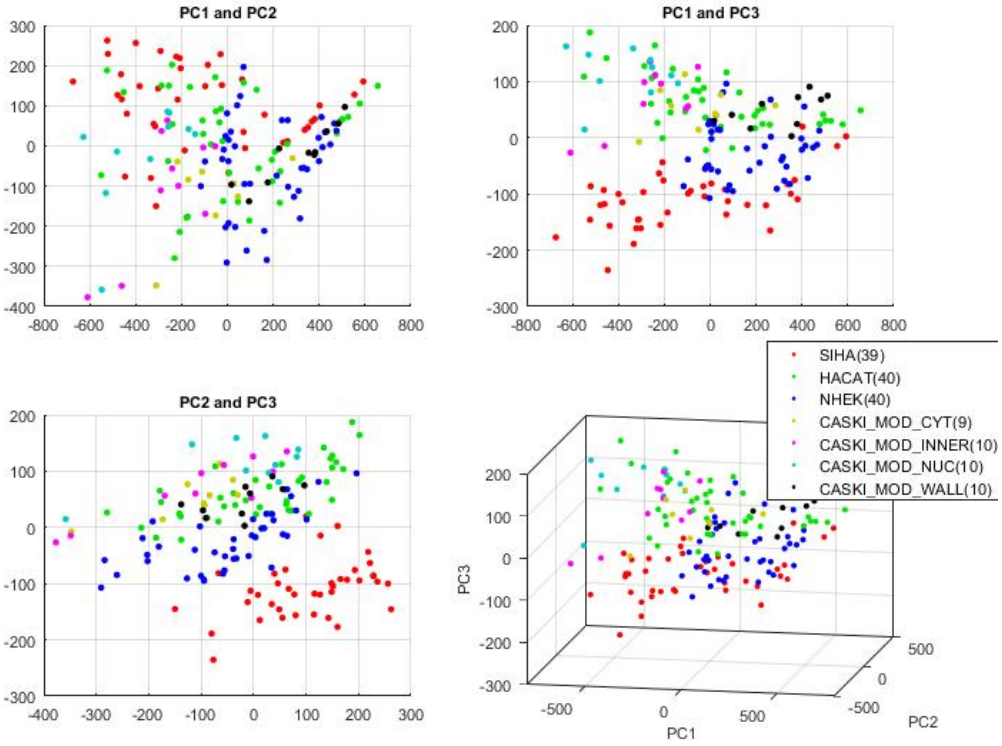


FIGURE 3.16: CaSki modulated PCA plot expanded by intracellular location.

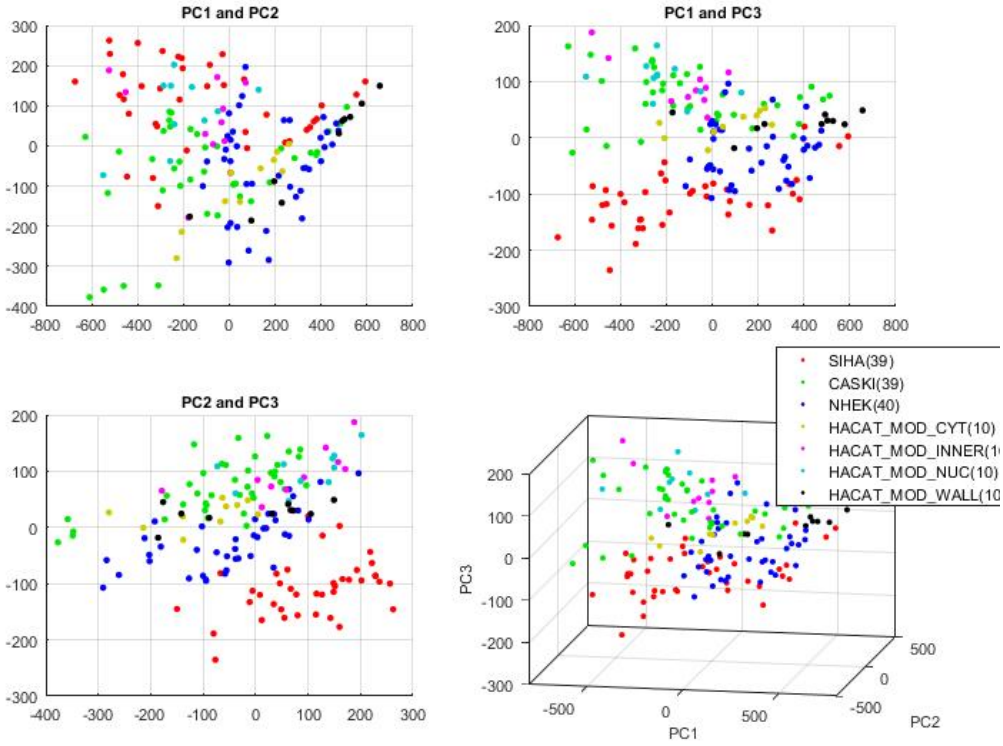


FIGURE 3.17: HaCaT modulated PCA plot expanded by intracellular location.

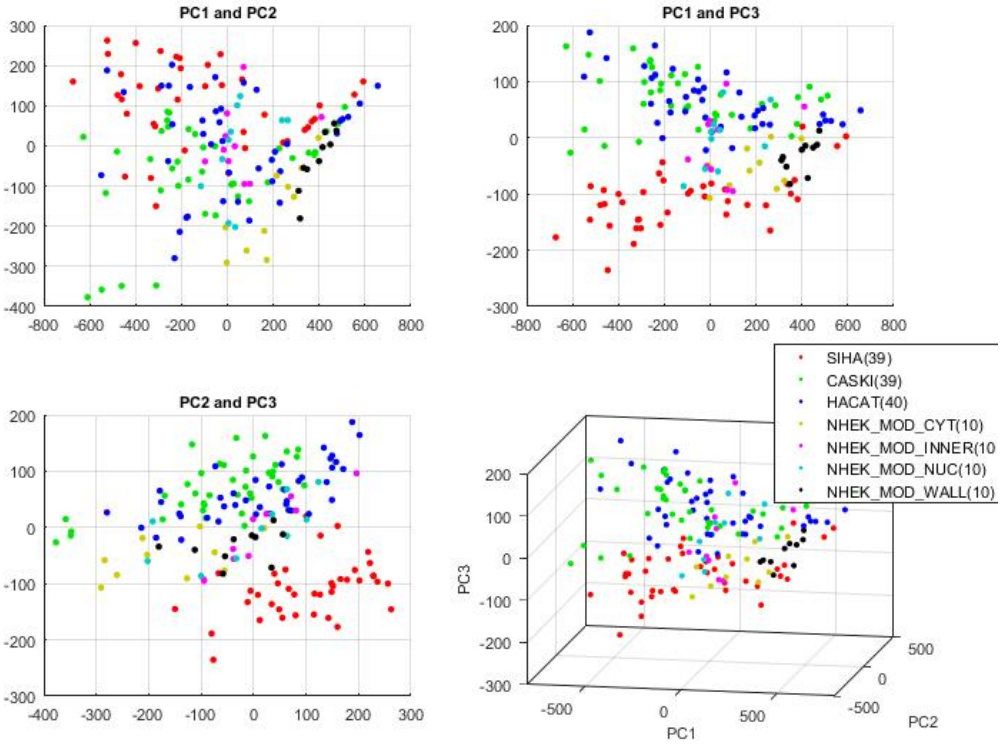


FIGURE 3.18: NHEK modulated PCA plot expanded by intracellular location.

3.7 Spectral variation due to sample storage

Finally, the potential effects of storage time on spectral classification were investigated. This is an important consideration for cervical smears as samples are stored indefinitely for archive in PreservCyt and may have to be revisited for clinical trials or in the case of ambiguous diagnoses.

Work by Hobro et al. considered spectral differences between live mouse embryonic fibroblast (MEF) cells and those fixed with different agents, including methanol [144]. Methanol acts as both a fixing and permeabilisation agent which dehydrates cells and precipitates proteins. In that case, a significant loss of overall Raman signal was observed in cells fixed with methanol, which were mostly associated with lipid bands. The effects of storage in fixative were not considered by Hobro et al. Based on the routine morphological and genomic analysis of stored smear samples without alterations to data, we hypothesised that changes over time after initial fixation would not be significant. If such changes do occur, it would have considerable impact on the use of Raman spectroscopy as a tool for clinical sample classification where samples may be archived in fixing solution for long periods of time either before analysis, for revisiting borderline samples or for comparing with new samples for disease progression.

A total of 150 cells were interrogated from SiHa cell samples which had been stored at $-20\,^{\circ}\mathrm{C}$ for three days, one month or six months in PreservCyt before following the same protocols described previously. The average Raman spectra are in Figure 3.19 - variation can be seen in the 1200-1500cm^{-1} region as well as for the Amide I (α-helix) peak at 1680cm^{-1}. Figure 3.20 shows the corresponding PCA plot, where grouping can be seen for the three day and six month clusters in the PC1/PC2 and PC2/PC3 projections, but not for the one month cluster. The confusion matrix (Table 3.7) shows considerable variation in classification accuracy, which indicates that there are sufficient spectral differences to 'classify' cells from the samples used. However, there is no trend according to storage time, which indicates that other sources of variation are present. It is posited that the variation seen may be due to different subculturing ages of the established cell line samples, which would not be an issue for clinical samples. However, the sensitivity of the technique here highlights the potential challenges that may arise from non-specific variation of cervical smears.

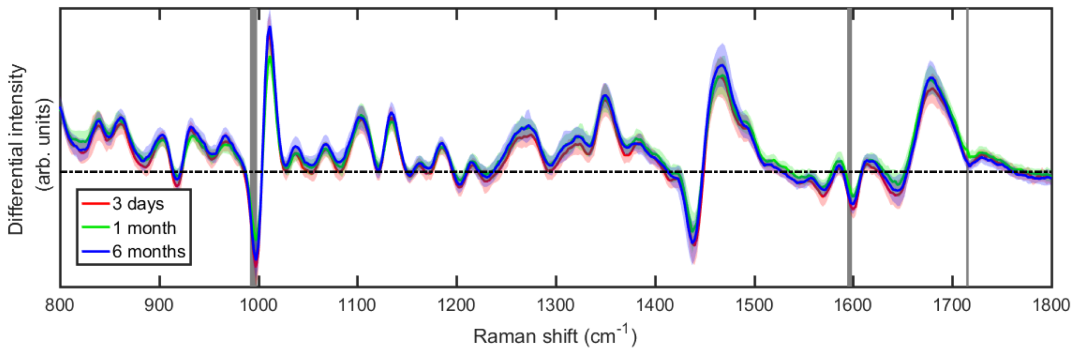


FIGURE 3.19: Variation of average Raman spectra of fixed SiHa cells with storage time after fixing. Cells were stored in PreservCyt at -20°C.

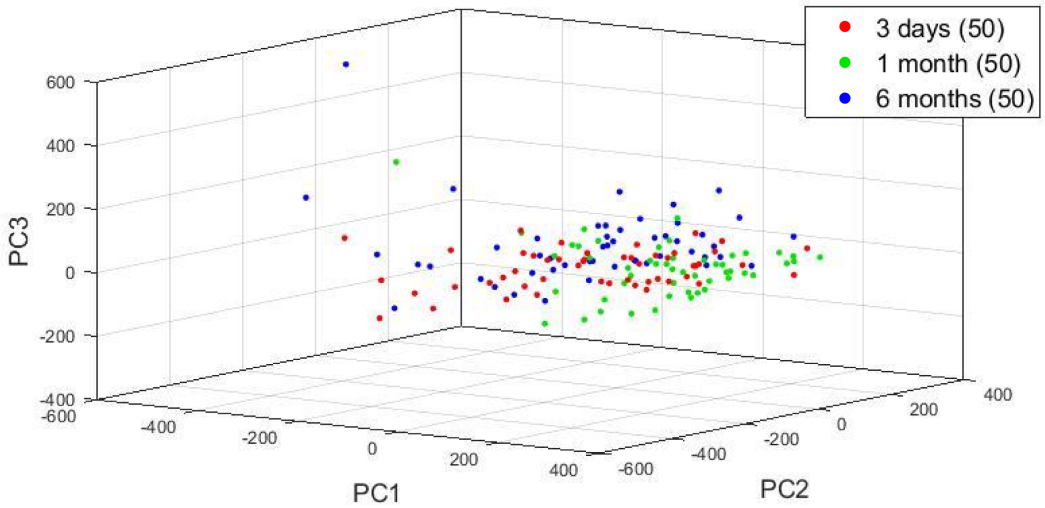


FIGURE 3.20

	Predicted			
Actual	3 day	1 month	6 months	Class acc.
3 day	**29**	0	21	58.0
1 month	1	**47**	2	94.0
6 months	17	1	**32**	64.0

TABLE 3.7: The confusion matrix resulting from cross-validation of the fixed cell storage time dataset, with accompanying percentage classification accuracy.

3.8 Analysis of significant vibrational modes

The direct interrogation of vibrational modes with this method can yield biologically relevant data regarding global changes to cell molecular populations. Although Raman is mode- rather than molecule-specific, conclusions may be drawn relating to alterations in the proteome, lipidome, epigenome and so on.

In Table 3.8, pairwise comparisons are made of the transformed cell lines versus NHEKs, in both the spontaneous and modulated acquisition modes. The wavenumber regions shown in bold refer to comparisons in which the NHEK spectrum is of higher intensity - this was not usually the case. Regions shown in italics refer to the statistically significant region being in the shoulder of the peak. Differences in shoulder position and intensity occur as the result of the molecular environment in which the vibrational bond is found.

As discussed in Section 3.5.2, the modulated differential-like analysis highlights different areas of the peak structure. The increased classification accuracy of WMRS may be due, in part, to the use of this spectral profile, which suggests that it may be a beneficial approach for standard acquisition.

The only regions which were statistically significant for all standard Raman spectral comparisons were those corresponding to Amide I: α-helix and Amide III: β-sheet, adenine and cytosine. It is posited that the increased density of protein conformational changes may arise from histones[117]. Many of the remaining significant peak differences also result from DNA-associated structures, including DNA O-P-O backbone stretching and polynucleotide chains.

The only differential region which demonstrated statistically significant variations in each case was that of the symmetric ring breathing mode of phenylalanine/tryptophan. There was also statistically significant reduction in the CH_2 deformation in lipids for CaSki and HaCaT cell modulated spectra compared to NHEKs. In the modulated case, the HaCaT cells also demonstrated a significant reduction in the ring breathing in tyrosine, and the C-C stretching modes in proline, but an increase in Amide I and III. Interestingly, there were no modes in the modulated case that distinguished the spontaneously transformed line (HaCaT) from the two cervical cancer lines (SiHa and CaSki), though in the standard case they could be distinguished with the concentration of DNA bases.

Significant Raman spectral regions (cm^{-1})						
Vibrational mode	NHEK-SiHa (s)	NHEK-CaSki (s)	NHEK-HaCaT (s)	NHEK-SiHa (m)	NHEK-CaSki (m)	NHEK-HaCaT (m)
DNA O-P-O backbone stretching					**800**	
Ring breathing in tyrosine/C-C stretching in proline/DNA O-P-O backbone stretching					**837-842, 856-866**	***844***
C-C skeletal stretching in protein					**931-936**	**934**
Symmetric ring breathing mode of phenylalanine/tryptophan ring breathing		1004		*1016-1018*	999-1023	*1014-1021*
ν(C-N)/DNA O-P-O backbone stretching/C-C amino side chains		1092-1096	1101-1103		1103-1119	1110-1117
C-C skeletal stretching					1138-1145	
Cytosine, guanine, adenine/antisymmetric phosphate vibrations					*1192-1197*	
Amide III: α-helix & collagen, CH in phospholipids, C-C in fatty acids	1230-1264	1220-1274	1237-1267, *1269-1291*		1269	*1267-1269*
Polynucleotide chain (DNA bases)	1332-1354, *1362-1378*	1324-1371			*1352-1366*	
CH_2 deformation in lipids/adenine/cytosine/C-O symmetric stretch			*1328-1392*		1387-1401	
CH_2/CH_3 deformation of lipids & collagen/CH_2 wagging/ CH_2/CH_3 deformation			*1490-1495*		1448-1469 *1448-1469*	***1436***
Adenine/guanine			1596			
Amide I: β-sheet/C-O stretch/ν(C-C) trans, lipids, fatty acids	*1680-1692*	*1681-1683*	*1680-1686*		1685-1720, 1685-1720	1699-1709
C-O lipids					1725	

TABLE 3.8: Statistically significant Raman spectral regions for used for classification of cell lines. Vibrational modes are assigned based on previous literature (Jess et al. 2007, Talari et al. 2014). Regions shown in bold represent those for which the NHEK signature was higher than that of the comparator. Regions shown in italics refer to peak shoulders rather than the peak itself (shown in normal text).

3.9 Conclusions

In this Chapter, fluorescence-free wavelength modulated Raman spectroscopy has been used to consider the benefits and practicalities of such a system in the context of label-free classification of fixed squamous epithelial cells in suspension, such as those produced during a cervical smear test.

The use of wavelength modulation in Raman spectroscopy and the resulting removal of autofluorescence provides clear benefits to its ability to automatically classify cells over conventional Raman scattering, without the requirement for manual baseline fitting, additional lasers or optical gates. Direct comparison with standard Raman spectroscopy shows marked improvement of sensitivity and specificity when considering both HPV (sensitivity +12.0%, specificity +5.3%) and transformation status (sensitivity +10.3%, specificity +11.1%). WMRS is able to provide superior cell classification accuracy, not just overall but in every case except CaSki/NHEK sensitivity. The ability to discriminate between HPV positive and negative cervical carcinoma cell lines is an encouraging step towards the potential integration of HPV and smear testing using optical discrimination. However, it must be noted that these carcinoma lines represent different mechanisms of origin, and by extension molecular pathologies. The ability to also identify HPV status in normal cervical cells would be a requirement of an integrated platform and so this will require consideration.

The investigation of location impact for spectral sampling provided important data regarding the expected differences between sampling particular cell morphologies or methods of collection. It is suggested that the lower specificity of sampling location in this dataset compared to previous experiments is due to the superposition of spectral data from different components of the cell, as the surface area occupied by the cell is lower. This lower margin of error could provide benefits for coupling the spectrometer with automated cell sampling capabilities. Translation to clinical samples would require additional studies into sampling spot size and locational sensitivity to account for the morphological differences which arise from cell provenance and harvesting technique.

Finally, the cell spectral data does not appear to be directly affected by fixed cell storage time which is a positive result in the context of long-term archive classification. The data does suggest however that other factors involved, such as biochemical changes which occur with the level of cell confluency. If this is the case, such *in vitro* artefacts are unlikely to affect clinical samples. However, the sensitivity of the technique to these changes suggest that similar issues with sample heterogeneity will have to be addressed with large clinical datasets.

The major technical barrier to this technique becoming clinically relevant is the acquisition method, which must be progressed beyond point sampling in order to achieve feasible sampling times. It is entirely feasible that the technique could be combined with other imaging methods such as Optical Coherence Tomography (OCT) [145] or Digital Holographic Microscopy (DHM) [146] in order to provide a cell 'topography' map for automated point sampling. The use of a digital guidance system could highlight unusual morphologies, thereby reducing the number of cells assessed per sample. It is hoped that such a system could be introduced to clinical settings in order to reduce the burden on manual cytological diagnosis, as well as increasing the likelihood of a correct outcome for patients.

Chapter 4

Development of antibody-conjugated nanoparticles

4.1 Introduction

As discussed in Chapter 1 of this thesis, the introduction of noble metal nanostructures into Raman environments allows for signal amplifications up to the order of 10^9 [99] compared to standard inelastic scattering events, dependent on the substrate-analyte complex. The mechanism of this surface-enhanced Raman scattering (SERS) phenomenon is discussed in Section 4.2. As such, SERS represents a way of overcoming the issues of acquisition times discussed at the conclusion of Chapter 3 without the need for alternative optical setups. Indeed, silver nanoparticles have been used directly with healthy and cervical cancer patient blood samples to amplify the Raman spectra and carry out PCA-based classification [124].

The ability of plasmonic nanostructures, and in particular nanoparticles (NPs) to preferentially interact with Raman scattering events at close proximity means that they may be used as a specific label as well as a general amplifier, becoming a vehicle for functional molecules. In the context of cervical cancer diagnosis, NPs may be conjugated to antibodies specific for a known marker of disease such as p16, allowing for expression of this target to be assessed in biological systems. Label-free SERS, in which metal nanostructures are used for direct enhancement, is well documented [147] but not suitable for this project: biological moieties would be preferentially amplified in areas of greater uptake, which is likely to vary depending on the morphology and orientation of slide-based cytology samples.

The development of antibody-conjugated nanoparticles or 'nanotags' for SERS applications is multifaceted. It requires not only the optimisation of the individual assembly components as seen in Figure 4.1 - namely the substrate, Raman label, blocking

layer, conjugation method and binding ligand - but also consideration of the potentially complex interplay between these elements and their surrounding environment, which often may only be inferred from optical data due to the size of the system.

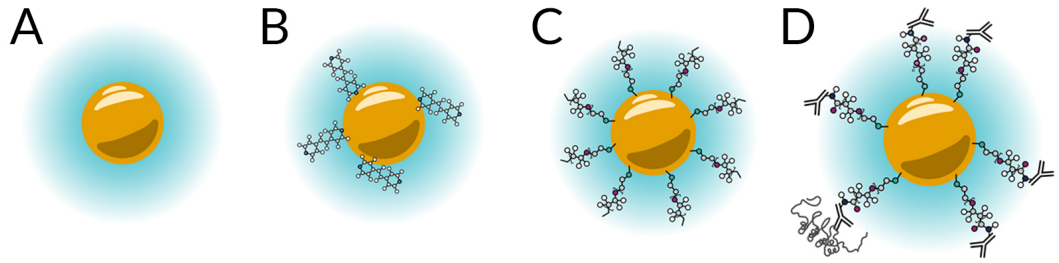


FIGURE 4.1: Schematic representations of A) bare gold nanoparticles (AuNPs) which absorb strongly in the green region of the visible electromagnetic spectrum and amplify local Raman signals. B) Reporter AuNPs or R-NPs which provide known Raman labelling. C) Blocked AuNPs which prevent non-specific interaction with the environment and improve *in vivo* circulation time. D) Targeted AuNPs which attach to molecules of interest via a binding ligand.

Nanoparticle substrates

SERS can be achieved with any nanoscale metallic structure as described in Section 4.2, and has also been shown with organic and inorganic semiconductors [148, 149, 150]. Noble metal nanoparticles (AuNPs and AgNPs) are most common owing to minimal oxidation and the ability to support electromagnetic enhancement in the visible and near-infrared regions, which are more relevant to biomedical imaging applications owing to their non-ionising effects. AuNPs demonstrate narrower wavelength tunability and less pronounced SERS enhancement owing to a lower Q-factor, but are more widely used in current clinical applications [151, 152]. Both types can be fabricated by a number of methods including metal salt reduction techniques [153, 154], etching and lithography [155, 156], laser ablation [157, 158] and biosynthesis [159, 160] amongst others.

Examples of colloidal solid nanostructures can be seen in Figure 4.2 and include spheres, rods, plates, wires, stars and pyramids amongst others [161]. More complex structures may be hollow [162], contain cores of different materials [163] or be deliberately aggregated within a solid dielectric [164]. Solid plasmonic substrates were not appropriate to this project but are discussed elsewhere [165].

The wavelength of laser excitation of localised surface plasmon resonances (LSPRs) should be close to the resonant peak of the substrate, which changes with size, shape

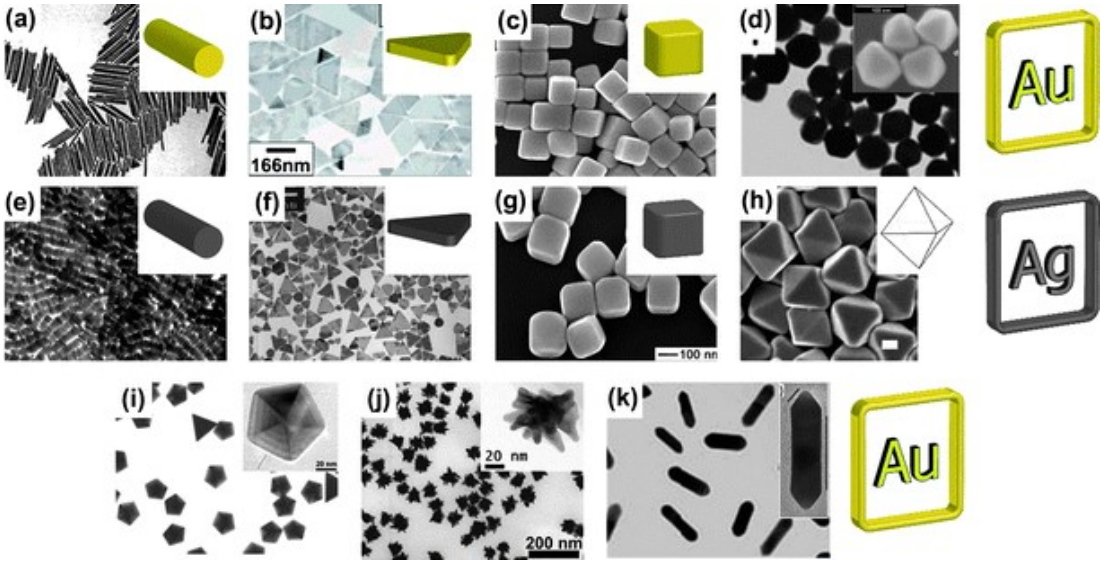


FIGURE 4.2: AuNPs with (a) rod-like, (b) triangular, (c) cubic and (d) octahedral shape. Ag NPs with (e) rod-like, (f) triangular, (g) cubic and (h) octahedral shape. AuNPs fashioned as (i) Au decahedra, (j) nanostars and (k) rods with sharp tips. Reproduced from Romo-Herrera et al. 2011 [161].

and material (Figure 4.3). Theory and experiment indicate that the greatest plasmonic enhancement is achieved when the nanoparticle resonance wavelength (λ_{NP}) is equal to the average wavelengths of the laser excitation (λ_{exc}) and Raman vibrational peak of interest (λ_{RS}), such that $\lambda_{NP} = 1/2(\lambda_{exc} + \lambda_{RS})$ [166].

Raman labels

In principle, any suitably soluble Raman-active molecule can be adsorbed onto a plasmonic substrate for electromagnetic enhancement. In practice Raman labels or 'reporters' fall into three groups as seen in Table 4.1: small molecules, bio-orthogonal molecules and fluorescent dyes. Small Raman-active molecules such as 1,2-di(4-pyridyl)ethylene (BPE) or 4-mercaptobenzoic acid (MBA) can be adsorbed or covalently bound by high-affinity sulphur (thiol) groups onto metallic surfaces with little impact on the hydrodynamic diameter of the structure. Bio-orthogonal molecules [168] such as alkynes, nitriles and stable isotopes of carbon and hydrogen produce peaks in the cellular silent region, reducing the need for complex signal processing. Fluorescent dyes with Raman activity are often used in development work so that activity can be confirmed through multiple imaging modalities. Small Raman molecules have the dual benefit

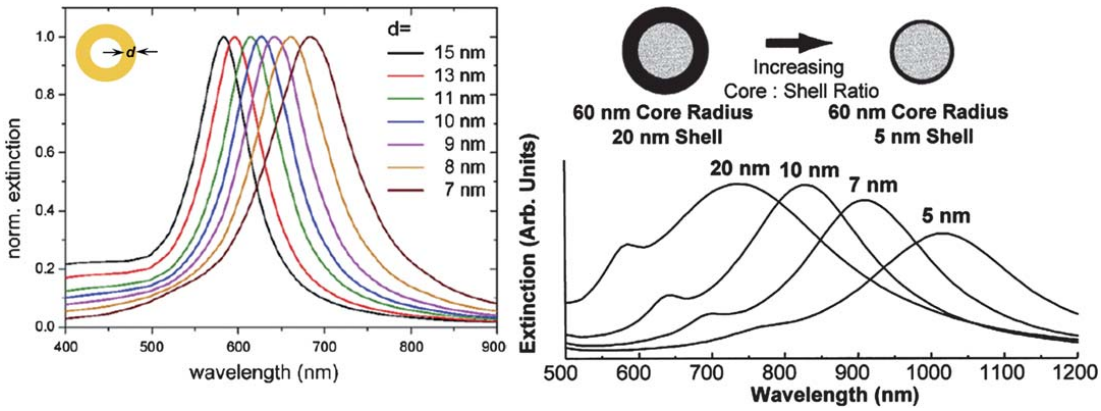


FIGURE 4.3: (Left) Calculated extinction spectra of single Au/Ag nanoshells with a radius of 27.5nm and varying shell thickness (d). (Right) Calculated extinction spectra of silica core/gold shells NPs (core radius 60 nm) with varying shell thickness. Reproduced from Wang and Schlucker 2013 [167]

of ease of spectral processing due to a lower number of vibrational peaks as well as their size.

Blocking layers

Additional coatings may protect the nanoparticle from adverse interactions with the surrounding environment which may lead to uncontrolled aggregation. This can happen as the result of changes in pH or opsonisation, where biomolecules (typically charged proteins) in the surrounding environment form unwanted coronae on the substrate [169]. Common protective layers include passivating proteins such as bovine serum albumin (BSA), silica (SiO_2) shells [170], polymers such as polyethylene glycol (PEG), glycans and lipid bilayers [171]. Polymer layers can also be used as spacers or modified with functional groups for conjugation as seen in Section 4.6. Spacer length is typically optimised so that both steric hindrance of functional molecules and overall particle size are minimal.

Binding ligands

Conjugated nanoparticles have been applied to a vast array of biosensing and biomedical applications as described previously, with target binding ligands including complementary DNA [172, 173, 174], siRNA [175], naturally occurring protein complexes including immunoglobulins [176], antibody fragments [177, 178] and synthetic peptides [169].

Reporter type	Examples
Small polarisable molecules	**Thiol groups**; MBA, MP, BT, BPT, ATP. **Adsorbing molecules**; BPE, PPY, PYOT, AZPY.
Bio-orthogonal molecules	**Alkynes**; phenylacetylene, propargyl glycine. **Diynes**; polyynes ('carbon atom wire'). **Nitriles**; MBN, phenylthiocyanate. **Stable isotopes**; carbon-13, deuterium.
Fluorescent dyes	**Fluorone dyes**; Rhodamine 6G, fluorescein. **Isothiocyanates**; FITC, MGITC and TRITC. **Cyanine dyes**; DTTC iodide (Cy7) **Triarylmethane dyes**; Crystal Violet, Brilliant Blue

TABLE 4.1: Raman-active molecules used as nanoparticle labels, grouped by molecule type. See Abbreviations (1) for full names.

For this project, quantification of p16 expression was carried out with commercially produced monoclonal antibodies. Antibody-conjugated nanoparticles have been used previously for cell separation [179] although these use cell surface markers which are rarely disease-specific. Such systems may be used for separating cell types in heterogeneous clinical samples. They have found use as EGFR-specific contrast agents in MRI imaging of oesophageal cancer [180], though this approach suffered from low sensitivity.

Gold nanoparticles conjugated with antibodies have been used extensively with a number of imaging techniques [181] including multiphoton plasmon resonance spectroscopy [182], optical coherence tomography [183] and surface-enhanced Raman spectroscopy, as well as through the simple use of the light scattering properties of AuNPs incubated with cells [184, 185].

The National Science and Technology Development Agency, Thailand, have produced AuNPs passively conjugated with both monoclonal and polyclonal antibodies to p16 for use in electrochemical [186] and paper-based immunoassays [187] respectively. At the time of writing only one other SERS-active conjugate to p16 has been produced [125], using a functional PEG linker, squaraine dye and p16/Ki67 monoclonal antibody (mAb) mix. It is not known how the levels of these two common stains for cervical intraepithelial tissue were quantified given the resulting heterogeneous NP population.

Conjugation methods

Interactions between antibodies and gold nanoparticles may be passive or supported with additional binding ligands. In the passive case, antibodies interact with the gold surface through hydrophobic or electrostatic attraction, or through co-ordinate (dative covalent) binding of sulphur atoms in the antibody to the gold surface.

Functional linkers may be used to mediate covalent binding of antibodies to colloidal gold. One of the most popular general methods is the use of a thiolated polymer, typically polyethylene glycol (PEG), which adheres to the metal surface with an additional functional group at the opposing end. The length of the polymeric spacer can be varied depending on use and also acts to sterically block non-specific interactions with the charged surface. The functional group can take multiple forms including carboxylate, aldehyde, hydrazide, thiol, hydroxyl and amide groups [188].

Biological linkers may also be used for antibody conjugation. Unmodified antibodies can be attached via Protein A or G presented on the surface of *Staphylococcus aureus* and *Streptococcal* bacteria respectively [189]. They both interact with the heavy chains of antibodies (typically the Fc region) and have varying binding affinities for the Fc portion of mammalian immunoglobulins. Native Protein G contains an albumin-binding site which is removed in the recombinant form to prevent non-specific binding in BSA-blocked environments. Finally, streptavidin-coated AuNPs may be used to immobilise biotinylated antibodies. This requires careful titration of the degree of biotin labelling so that antibodies do not bind to multiple streptavidin-coated NPs [190].

The three methods investigated during this project - passive, covalent and bioconjugation - are summarised schematically in Figure 4.4 and discussed in greater detail in Sections 4.6.1 , 4.7.1 and 4.8.2 respectively.

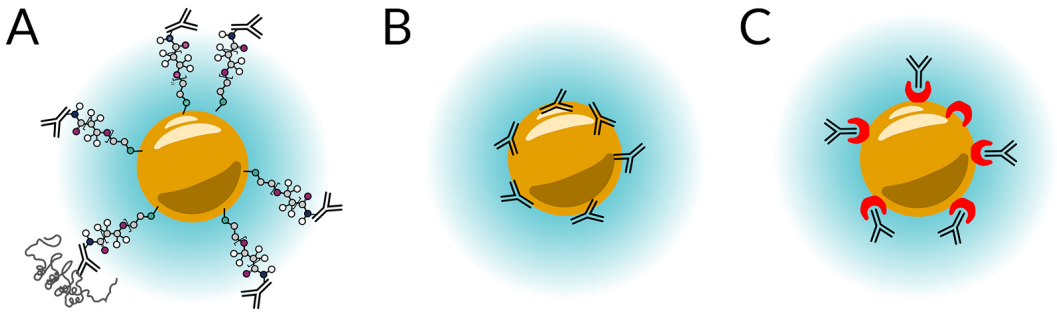


FIGURE 4.4: Schematic of the conjugation processes in A) covalent conjugation using a functionalised PEG linker with carboxyl group, B) passive conjugation and C) bioconjugation using Protein G' (shown in red).

4.2 Concept: SERS and nanoparticle conjugates

As discussed in the Introduction, spontaneous Raman scattering is an intrinsically rare occurrence which takes place as the result of anisotropic alterations in the polarisation of a molecule with the electric field of an incident photon. In the phenomenon of surface-enhanced Raman scattering by electromagnetic enhancement, this molecular polarisation is amplified by the presence of additional electric fields from surface plasmons. Surface plasmons are the collective oscillations of conduction electrons, excited by electromagnetic radiation at a metal-dielectric junction.

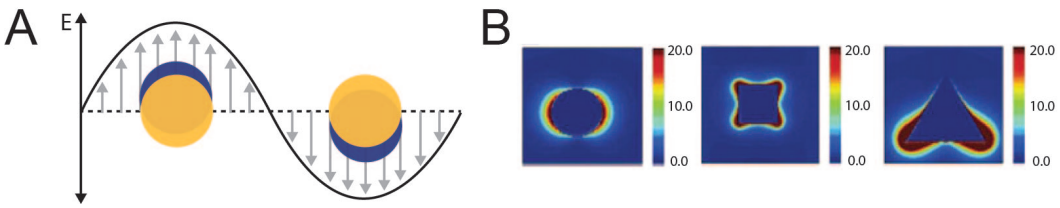


FIGURE 4.5: A) Schematic of localised surface plasmons induced by the electric field E of an incident laser. The free electron cloud (blue) of the gold sphere (yellow) is subsequently displaced, generating a localised electric field, simulated in B) as E^2 for a sphere, cube and triangular pyramid (reproduced from Kahraman et al. 2017) [191].

If the metal interface is continuous, plasmonic oscillations propagate between 10 and 300nm as surface plasmon polaritons (SPPs): on a surface such as a sphere or at sharp edges, the oscillations are limited to localised surface plasmon resonances (LSPRs) as shown in Figure 4.5. Additional chemical enhancement may occur as the result of charge transfer between the metal substrate and certain adsorbed molecules,

and coupling with LSPRs from neighbouring structures can result in 'hotspots' at particle junctions[192]. The magnitude of the electric field at the surface of a spherical nanostructure, E, has the following proportionality with the incident electric field E_0

$$E^2 \propto E_0^2 |\frac{\varepsilon_m - \varepsilon_0}{\varepsilon_m - 2\varepsilon_0}|^2 \tag{4.1}$$

where ε_m is the wavelength-dependent dielectric constant of the metal surface, and ε_0 is the dielectric constant of the local environment around the sphere [193].

Nanostructures can be fabricated which support LSPRs and therefore act as Raman scattering amplifiers at close range (typically 0-4nm). The SERS enhancement factor is directly proportional to the fourth power of the plasmons' electric field intensity, with the magnitude of supported plasmons directly proportional to a characteristic quantity of the nanostructure known as the quality factor, Q.

$Q = \omega \frac{\partial \varepsilon_r}{\partial \omega} / 2(\varepsilon_i)^2$, where ω is the excitation frequency of the laser, and ε_r and ε_i are the real and imaginary parts of the structure's dielectric function respectively [191]. Q depends then on the size, shape and material of the nanostructure but also on the dielectric used, corresponding to the solvent or blocking agent for colloidal systems. Experimentally, the average substrate enhancement factor (EF) can also be described as

$$EF = \frac{I_{SERS}/N_{surface}}{I_{RS}/N_{volume}} \tag{4.2}$$

where I_{RS} is the Raman intensity produced by the average number N_{volume} of molecules within the scattering volume and I_{SERS} is the SERS intensity of the average number $N_{surface}$ of molecules adsorbed onto the substrate within the same volume [194].

For this project, spherical gold nanoparticles were chosen as the substrate owing to ease of production, plasmonic anisotropy and lower toxicity and aggregation probability than silver NPs in the instance that they be transferred to topical *in vivo* studies. In practice, gold nanoparticles ranging from 5-150nm can be used for SERS applications: any larger and the optical enhancement deteriorates due to decreasing surface/volume ratio. Choice of diameter involves balancing optical enhancement favoured by smaller particles and diagnostic sensitivity supported by particles with larger surface areas, dependent on the application. A nanoparticle diameter of 40nm was chosen as the most common size for potential lateral flow applications, representing a compromise between optical contrast and antibody coverage. Small Raman

molecules were chosen for their stability and minimal disruption to nanoparticle diameter, which was an important consideration for the use of conjugates in fixed cells. Commercially available antibodies were chosen as well-characterised binding ligands - the conjugation method is explored in the remainder of this chapter.

4.3 Materials and methods

All reagents were purchased from Sigma Aldrich unless otherwise stated. Ultrapure water ('dH_2O', MilliQ) was used for all experiments. All nanoparticle centrifugation ('pelleting') steps were carried out at 3400g for 20 minutes in the first instance unless otherwise stated. All shaking steps were carried out using a benchtop plate shaker (IKA MS 3) at 500rpm. It should be noted that as this chapter consists largely of method optimisation work, the protocols outlined below represent an initial starting point. Modifications are outlined in the main text.

4.3.1 Nanoparticle production

A 1l Erlenmeyer flask and magnetic stirrer were cleaned with aqua regia (HNO_3+3HCl) for two hours, rinsed and dried. 60.5mg sodium tetrachloroaurate (III) dihydrate was dissolved in 500ml dH_2O (equivalent to 304.17µM solution) and heated at 100°C and stable stirring conditions on a magnetic hotplate until boiling. 57.5mg sodium citrate tribasic dihydrate dissolved in 7.5ml dH_2O (equivalent to 26.07mM solution) was added and mixed for 15 minutes. This was left to cool with continuous stirring before transferring to a clean glass container.

4.3.2 Raman spectroscopy

Raman spectra of nanoparticle samples in bulk solution were made using a Sierra Series desktop Raman spectrometer (Snowy Range Instruments) with 100mW 785nm laser at full power. All spectral acquisitions were made for one second unless otherwise stated, with an undiluted 500µl sample stored in a cylindrical glass vial cleaned using aque regia. When direct sample comparison was required between measurement batches, spectra were normalised with reference to the characteristic 884cm^{-1} peak of ethanol to account for small fluctuations in laser power over time, with a normalising multiplication factor of $10000/intensity_{EtOHpeak}$.

4.3.3 Covalent conjugation

A method diagram for both approaches is shown in Figure 4.6.

On-NP method

In the standard protocol, 40µl 1mM carboxy-PEG_{12}-thiol ('CT-PEG', ThermoFisher Scientific) was added to 960µl reporter-AuNPs and shaken for 2 hours. Samples were pelleted and resuspended in 74µl EDC (1mg/ml), 217µl NHS (1mg/ml), 709µl 10mM MES buffer (pH6.0), and shaken for 30 minutes. Following centrifugation, the pellet was resuspended in 980µl 10mM HEPES buffer (pH7.0) and 20µl antibody or blank control (dH_2O) and shaken overnight. The completed conjugates were pelleted and resuspended in 980µl 10mM phosphate buffer (pH7.6) unless otherwise stated.

Off-NP method

In the standard protocol, 40µl 1mM CT-PEG and 74µl EDC (1mg/ml) were added to one 1.5ml Eppendorf, and 217µl NHS (1mg/ml) and 20µl antibody or blank control (dH_2O) to another. Both were shaken for 5 minutes before combining and adding 709µl 10mM HEPES buffer (pH7.0). This solution was shaken overnight, added to 1ml reporter-AuNPs and left to shake again for 2.5 hours. The completed conjugates were pelleted and resuspended in 980µl 10mM phosphate buffer (pH7.6) unless otherwise stated.

4.3.4 Passive conjugation

In the standard protocol, Raman reporters were adsorbed onto bare colloid by shaking 10µl appropriately diluted reporter solution with 990µl AuNPs for 30 minutes. Following centrifugation for 15 minutes at 2300g and resuspension in 980µl phosphate buffer (pH7.6), antibody solution (1-30,000ng/µl) was added to 1ml reporter-NPs and shaken for 15 minutes. 10µl 10%[w/v] BSA solution was added, shaken for 30 minutes and the solution pelleted and resuspended in 10mM phosphate buffer (pH7.6) unless otherwise stated.

4.3.5 Bioconjugation

Raman reporters were adsorbed onto bare colloid by shaking 10µl appropriately diluted reporter solution with 990µl AuNPs for 30 minutes. Following centrifugation for 15 minutes at 2300g and resuspension in 980µl MES buffer (10mM, pH4.8), 100µl

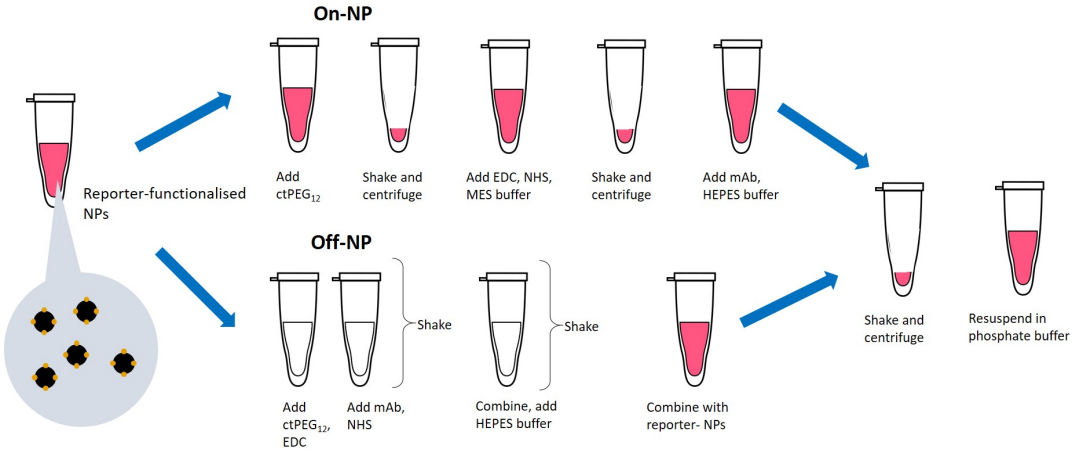


FIGURE 4.6: The two approaches used for covalent conjugation of antibody using EDC and sulfo-NHS linker chemistry. In the 'on-NP' method, the moieties are added in order to the nanoparticles with centrifugation steps in between to remove excess reagent. In the off-NP method, the carboxylated PEG and EDC, and NHS and mAb, are assembled in isolation. The ligand complex is then formed in an overnight step before it is added to the NP surface.

Protein G in 10mM phosphate buffer (pH7.6) at the required concentration was added and the solution left to shake for 30 minutes. The resulting solution was pelleted again, resuspended in 980µl MES buffer, added to 10µl mAb solution (mAb diluted 1:10 in PBS) and left to shake for 30 minutes. The completed conjugate was pelleted and resuspended in 980µl 10mM phosphate buffer (pH7.6).

4.4 Production of Raman-active nanoparticles

4.4.1 Gold nanoparticle characterisation

Gold nanoparticles were produced according to a modified Turkevich method, described in Section 4.3.1. The citrate reduction agent also caps the particles, producing a negative surface potential which allows them to remain in suspension. These colloids, nominally spherical and 40nm in size, were characterised using dynamic light scattering (DLS) and zeta potential measurements, extinction spectroscopy and transmission electron microscopy (TEM). Two batches of nanoparticles were produced during the project, with some variation in physical properties. The average diameter in nanometres of the nanoparticles for both batches was calculated from TEM measurements to give 48 ± 9nm and 41 ± 9nm for the first and second batch respectively (Figure 4.7).

From this the particle number density and concentration were generated using

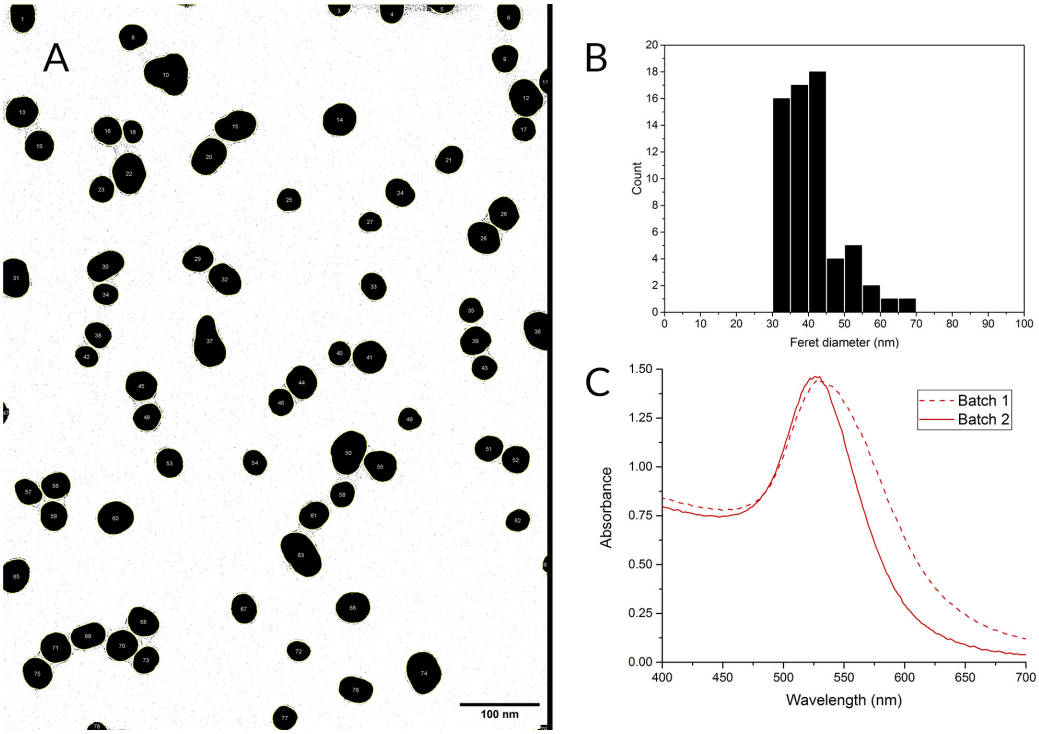


FIGURE 4.7: A; TEM image of nanoparticles, taken using a Tecnai TF20 cryo-TEM (FEI) and thresholded in ImageJ (FIJI) to provide a mask for particle sizing; B) population histogram of the Feret diameter of the second batch of NPs, demonstrating a modal diameter between 40-45nm as expected, and an average of 41 $\pm$ 9nm; C) the variation in absorbance profile between the first and second batch of nanoparticles used for the project.

$$N = \frac{A_{450} \times 10^{14}}{d^2[-0.295 + 1.36\exp(-(\frac{d-96.8}{78.2})^2)]} \tag{4.3}$$

where A_{450} is the absorbance at 450nm, and d is the particle diameter in nanometres, to give the values summarised in Table 4.2 [195].

The gold nanoparticles (AuNPs) had a localised surface plasmon resonance (LSPR) peak of 532nm and 526nm for Batch 1 and 2 respectively, with a full width half maximum (FWHM) of 60nm and 46nm for the same. Monodispersed solutions of AuNPs typically have a FWHM of less than 100nm, with narrower peaks corresponding to higher monodispersity.

4.4.2 Investigation of Raman reporters

Raman reporters, or strongly Raman active small molecules, were chosen for investigation over Raman-active fluorescent dyes in order to reduce the overall size of the

Specification	Batch 1	Batch 2
Number density (particles per ml)	5.38×10^{10}	7.39×10^{10}
Molar concentration (nM)	8.9	12.3
Feret diameter by TEM (nm)	48 ± 9	41 ± 9
Hydrodynamic diameter in dH_2O (nm)	62.8 ± 1.9	48.0 ± 0.3
Zeta potential (-mV)	37.7 ± 1.1	37.4 ± 0.9
LSPR peak (nm)	532	526

TABLE 4.2: Characterisation of the two gold nanoparticle batches produced. Errors correspond to a single standard deviation from the mean.

Raman-active conjugate. This is an important consideration for fixed cell incubation, as cellular uptake mechanisms are limited to passive diffusion and electrostatic interactions.

Initially BPE, PPY, PYOT, MBA and AZPY were investigated (see Figure 4.8 for details). This was done by diluting a stock concentration of reporter (typically 10mM in ethanol) in dH2O before adding 5µl diluted reporter to 495µl bare AuNPs (or equivalent ratio at larger volumes) and shaking for 30 minutes. The resulting reporter NPs were microcentrifuged at 3400g for 20 minutes, 480µl supernatant removed to eliminate non-adsorbed reporters, and the pellet resuspended in 480µl dH2O for characterisation.

Further experiments were then carried out with the most reproducible reporter-NP combinations. Serial dilutions of BPE, PPY and PYOT indicated that BPE provided the most stable coverage of AuNPs up to a micromolar limit, and so this was chosen for use with the targeted conjugate (Figure 4.9). Potential candidates for a physicochemically identical dummy particle were also considered (Table 4.3), although this strategy was revised in Sections 4.8 and 4.10 owing to the masking effects of antibody conjugation on the reporter layer.

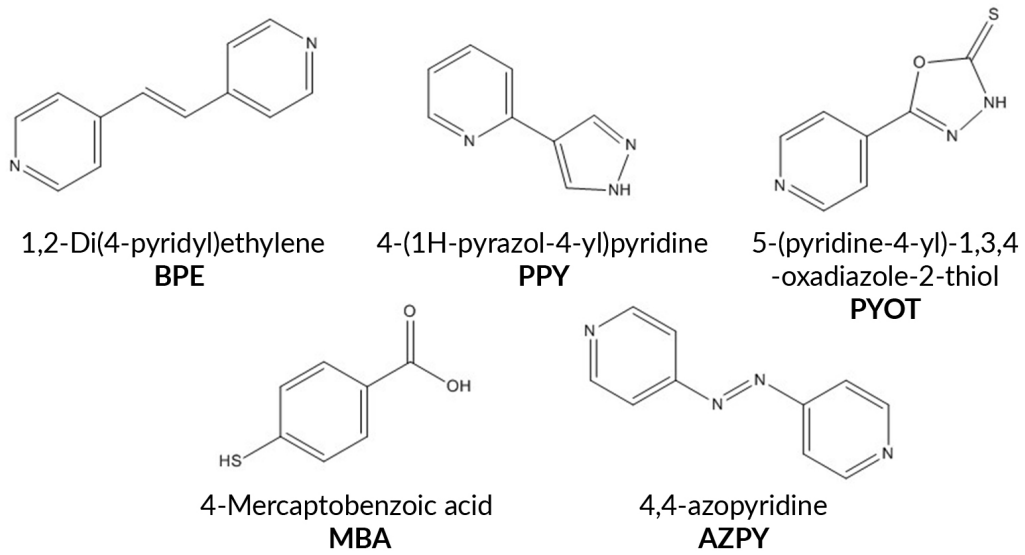


FIGURE 4.8: The structures, names and abbreviations of the five Raman-active molecules initially chosen for the project.

Sample	Hydrodynamic diameter (nm)	Zeta potential (-mV)
AuNP	62.8±1.9	37.7±1.1
BPE 10^{-5}M	62.9±2.8	9.07±7.1
BPE 10^{-6}M	62.8±2.8	37.3±0.8
BPE 10^{-7}M	54.4±1.5	31.0±6.8
PYOT $^{-5}$10M	57.2±0.6	40.5±4.4
PYOT $^{-6}$10M	54.9±0.4	24.8±5.5
PYOT $^{-7}$10M	55.0±0.5	27.4±3.2
PPY $^{-5}$10M	75.8±2.2	20.1±2.1
PPY $^{-6}$10M	59.4±0.7	31.0±3.5
PPY $^{-7}$10M	89.8±1.3	28.7±5.3

TABLE 4.3: Characterisation of the Raman reporter-coated particles produced at different initial concentrations. Zeta or surface potential is shown in negative electron volts, with higher values corresponding to greater solution stability due to electrostatic repulsion.

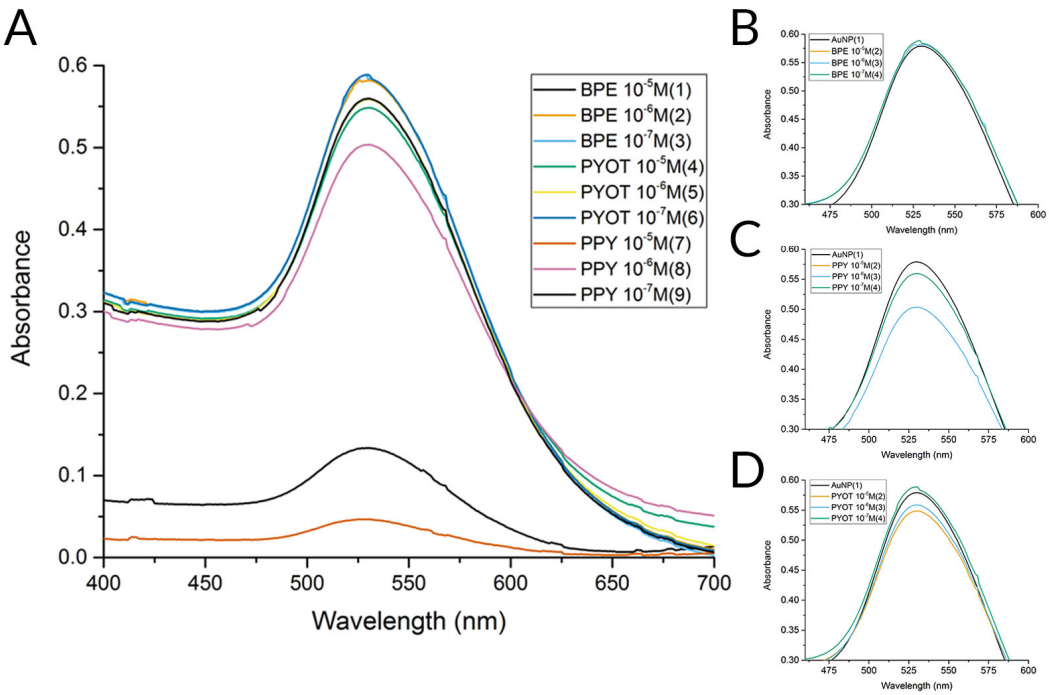


FIGURE 4.9: A) Absorbance spectra of the three chosen reporters indicating an order-of-magnitude concentration operating window for adsorption onto 40nm gold nanoparticles; B) BPE concentrations compared to bare AuNP absorbance spectra (10^{-5} M not seen due to low absorbance from aggregation); C) PPY concentrations compared to AuNP absorbance spectra (10^{-5} M not seen due to low absorbance from aggregation); D) PYOT concentrations compared to AuNP absorbance spectra.

4.4.3 Optimisation of BPE-AuNP stability and signal

More sensitive concentration experiments indicated that, under the centrifugation conditions outlined in Section 4.4.2 and when suspended in dH2O, an initial reporter concentration of 1.25µM was the maximum which could be stably achieved. Absorbance measurements and surface potential dropped rapidly at 1.50µM (Figure 4.10). Optimum surface charge (zeta) potential, possibly corresponding to monolayer coverage, was achieved at 0.75µM. Samples produced in triplicate supported this conclusion, with considerably less variation in optical characteristics compared to other samples (Section 4.7.4).

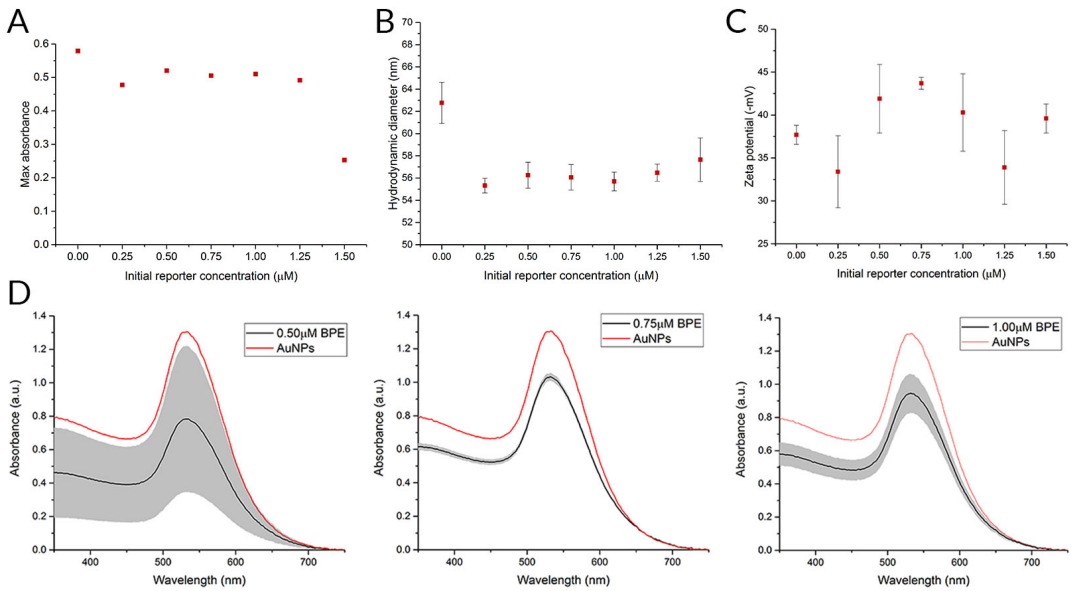


FIGURE 4.10: A) Absorbance peak of bare AuNPs and reporter-NPs between 0.50 and 1.50µM initial BPE concentration, showing an aggregation-induced reduction in absorbance at 1.50µM; B) variation in hydrodynamic diameter with addition of reporters, indicating a change in interaction of the colloidal structures with their solvent as a result of reporter addition; C) zeta potential of the same in three technical replicates, showing optimum charge stability of the colloidal solution at 0.75µM; D) the average absorbance of three sample replicates of initial BPE concentrations 0.50, 0.75 and 1.00µM, with a single standard deviation from the mean shown in grey. Again, this suggests that 0.75µM initial BPE concentration corresponds to monolayer coverage for 40nm AuNPs.

The Raman spectra for BPE at different concentrations can be seen in Figure 4.11. The major peaks are at 1019, 1202, 1340, 1608 (principal peak) and 1637cm^{-1}. Gaussian peak fitting of the resulting data following baseline removal indicated that a detection minimum of 4nM BPE in solution could be achieved under the conditions used above.

This was calculated using a linear extrapolation of each Raman peak at the concentrations produced until the axis intercept was reached (Figure 4.12).

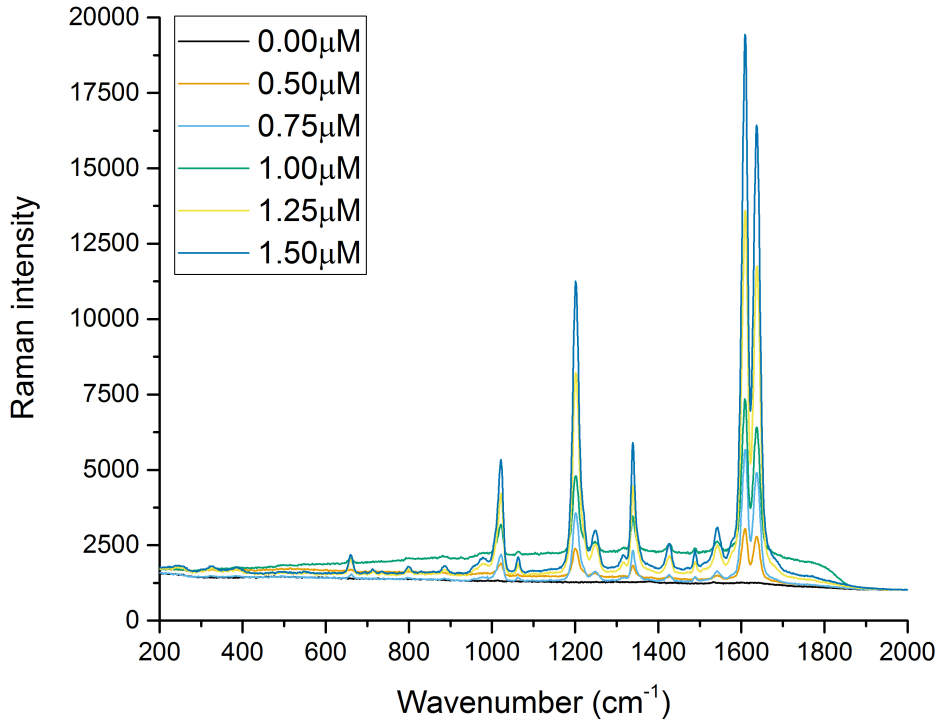


FIGURE 4.11: The fingerprint-region Raman spectra of bulk reporter-NP samples at initial BPE concentrations between 0.50 and 1.50µM, as well as bare AuNPs. Acquired over 1.00s with a 100mW 785nm laser.

Further optimisation of Raman signal from the SERS conjugates is considered in Section 4.8.4 once a conjugation method was chosen.

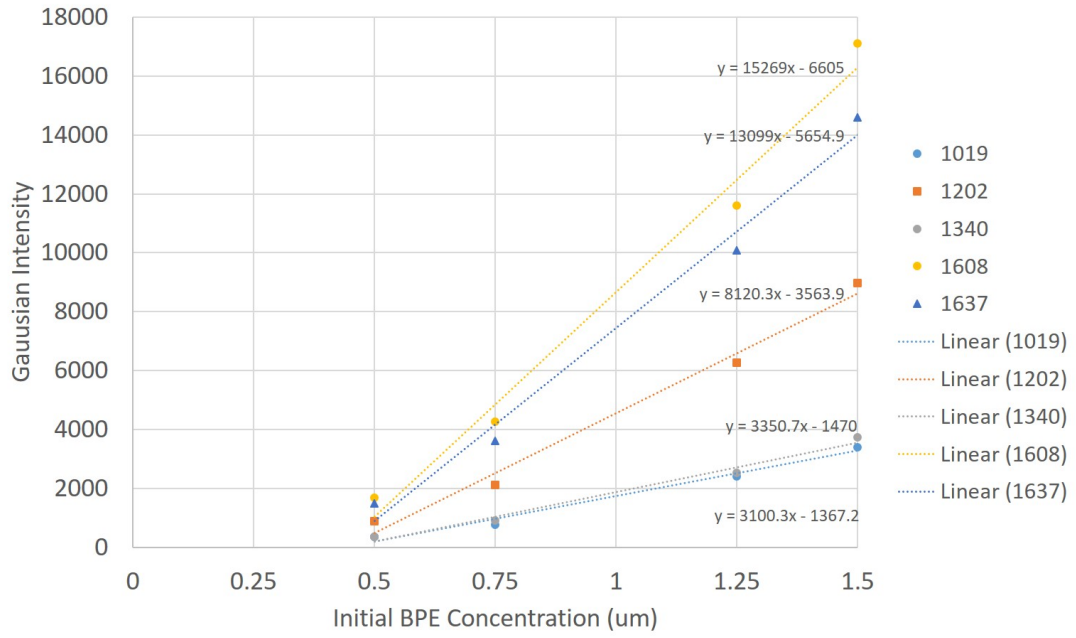


FIGURE 4.12: Intensity at each major Raman peak compared to initial BPE concentration. The linear relation between intensity and concentration for standard Raman spectroscopic measurements indicates that the spectral limit of detection (the point at which any of the primary peaks would be fitted with a zero-intensity Gaussian peak) is 0.40µM initial BPE concentration, equivalent to 4nm in solution with AuNPs.

4.4.4 Effect of buffers on signal and stability

Typically, reporter-NPs must be suspended in buffers other than distilled water so that subsequent reactions do not cause significant pH shift and particle aggregation. Reporter-NPs were prepared at three concentrations as described in Section 4.4.2 but with a final resuspension solution of one of the following:

- dH2O
- MES buffer (pH4.8, 10mM)
- Phosphate buffer (pH7.6, 10mM)
- HEPES buffer (pH7.0, 5mM)

MES (2-(N-morpholino)ethanesulfonic acid) is a buffer with minimal salt effects whose useful pH range is lower than that of other common biological solvents. It was used for bioconjugation experiments at pH4.8 (see Section 4.8) to passively adsorb Protein G onto the surface of the reporter-NPs. Phosphate buffer was used for passive

conjugation, as its high buffering capacity reduces likelihood of pH-dependent NP aggregation. HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) is used in the preparation of EDC-NHS antibody complexes in the covalent conjugation method.

The effect of buffers on the Raman signal of the nanoparticles can be seen in Figure 4.13. It is clear from this that a BPE-NP solution stored in phosphate buffer degrades considerably compared to others, and so should be avoided for conjugation protocols. This has not been seen in any published material but is posited to be due to differences in the dielectric constant of each buffer, which makes a contribution to SERS enhancement.

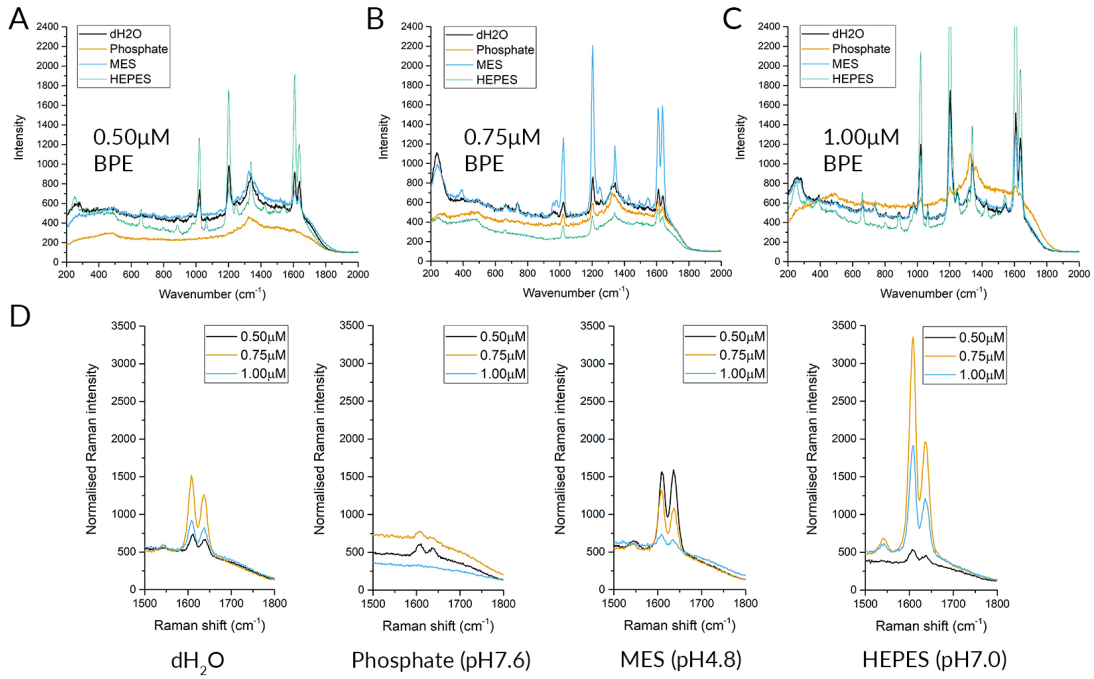


FIGURE 4.13: A comparison of Raman spectra of BPE-AuNPs at a range of concentrations in the buffers relevant to chosen conjugation chemistries. A) 0.50µM comparison; B) 0.75µM comparison; C) 1.00µM comparison; D) comparisons organised by buffer type, indicating a clear deterioration of Raman signal with phosphate buffer use.

It was unclear whether phosphate buffer was causing permanent dissociation of reporters from the surface, or if the effect was due to temporary interaction of buffers with the colloid. NPs stored in phosphate buffer for 30 minutes before centrifugation and resuspension in MES for overnight storage were found to maintain their signal, as were samples stored in MES briefly before phosphate buffer (Figure 4.14).

Further buffer conditions for storage of final conjugates are considered in Section

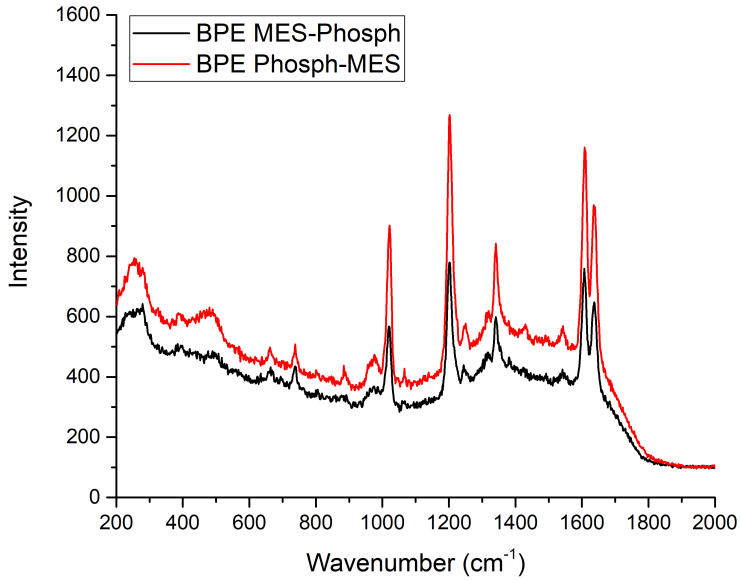


FIGURE 4.14: The Raman spectra of BPE-AuNPs (0.75µM) stored in MES then phosphate buffer (black line), and phosphate then MES buffer (red line). This was done to explore whether the spectral deterioration was a permanent or temporary effect.

4.9. Having synthesised a SERS-active AuNP substrate, the binding mechanism had to be characterised.

4.5 Antibody characterisation

4.5.1 Stability and sensitivity of commercial antibodies

Initially, a monoclonal antibody by BD Pharmingen (Purified Mouse Anti-Human p16, Clone G175-405) was used to assess p16 expression in the cell lines chosen for this project. The antibody is supplied as 31.25µg/ml in an aqueous buffered solution containing BSA, goat serum and sodium azide. Despite being advertised as routinely tested with Western Blotting, no publications were available. Western Blots for SiHa cells, which express p16 strongly, did not produce any signal when tested with antibody diluted to 1:1000, 1:500 and 1:200 in milk solution.

The antibody was then tested against purified p16 in a dot blot, where the protein was spotted down in either its native form as provided, or after denaturation by adding 20% [v/v] DTT and heating to 70°C as per the Western Blotting protocol

in Chapter 2. As seen in Figure 4.15, this antibody binds with much lower affinity to the denatured form of p16 which makes it much less versatile as a candidate for nanoparticle conjugation. Additionally, the supplied format is comparatively dilute - requiring much higher concentrations for a Western Blot than used previously - and is mixed with BSA which would compete for nanoparticle surface coverage.

Santa Cruz antibody (sc-9968), supplied as 0.2mg/ml in PBS with sodium azide and gelatin, was then tested and found to be suitable for Western Blotting (also Figure 4.15). This was used for conjugate development work as discussed in Section 4.6.3 before discovering that it was a discontinued product. As such, a newly released antibody from BD Pharmingen (Clone G175-1239) was tested and found to have a high affinity for p16 in Western Blot format, as well as being suitable for immunofluorescent work. All antibodies were diluted 1:1000 in milk solution for the blots shown in Figure 4.15. This is supplied as 0.5mg/ml in aqueous buffered solution containing sodium azide. The new G175-1239 clone was used for all subsequent experiments.

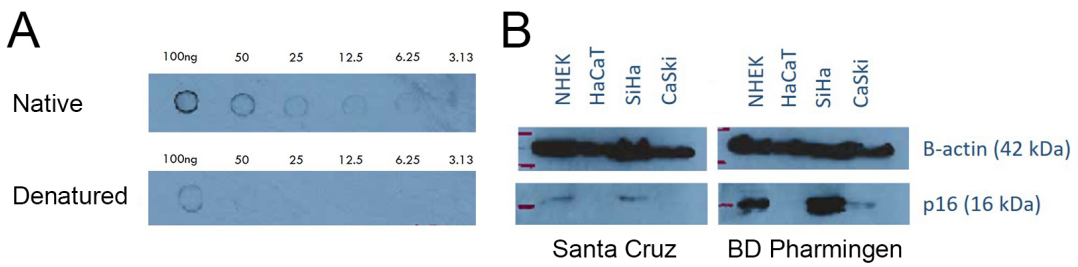


FIGURE 4.15: A) Dot blot of BD Pharmingen mAb against purified human recombinant p16 at different concentrations, native and denatured. B) Western blots with Santa Cruz antibody, now discontinued, and a newly available BD Pharmingen mAb. Note that the concentrations are different so a direct comparison cannot be made.

4.5.2 Antibody limit of detection

Following an initial test of the protocol (Chapter 2), the sensitivity of BD Pharmingen mAb Clone G175-1239 was tested in triplicate by ELISA at final concentrations from 2.000 to 0.031ng/ml in 1:1 serial dilutions with PBS. The lumimetric results can be seen in Figure 4.16, and are shown as a function of both p16 concentration and absolute protein content. Error bars are given as a single standard deviation from the mean value. Protein concentration values were fitted with a third-order polynomial to give the fit shown, with an adjusted coefficient of determination (COD, R^2) of 99.4%. The linear fit of absolute protein content yields a slightly poorer fit but can be used to

determine the limit of detection (LoD) of the antibody for this given format, which is a useful consideration in nanotag production to ensure sufficient binding sensitivity.

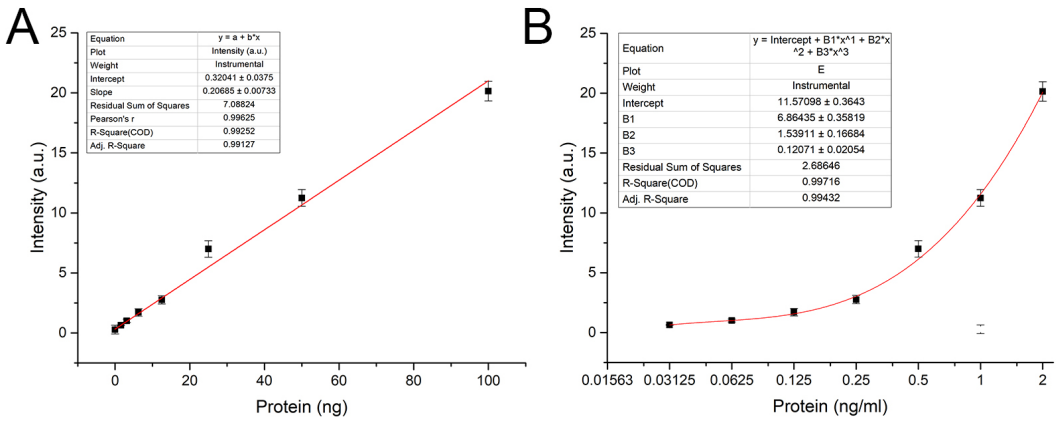


FIGURE 4.16: Lumimetric response of BD Pharmingen mAb Clone G175-1239 to purified p16 in an ELISA format, conducted in triplicate and shown in terms of protein concentration and absolute content.

The limit of blank (LoB), defined as the highest concentration of analyte which may appear as a blank signal[196], is given by

$$LoB = \mu_{blank} + 1.645\sigma_{blank} \tag{4.4}$$

for a Gaussian distribution with 90% confidence interval where 'blank' is the lumimetric value for N=8 empty wells with $\mu = 0.0108$ and $\sigma = 0.0104$. Thus, LoB = 0.028 by luminosity, which is negligible by protein concentration. The limit of detection, LoD, is given by

$$LoD = LoB + 1.645\sigma_{samples} \tag{4.5}$$

where $\sigma_{samples}$ was calculated as an average of the sample standards deviations, equivalent to 0.4209 luminosity. This gives LoD = 0.692 luminosity, equivalent to 1.80 ± 0.25ng or an initial concentration of 36 ± 5ng/µl, where the error was calculated as the maximum deviation from the linear fit.

4.6 Covalent conjugation of antibody

4.6.1 Covalent conjugation scheme

Following primary optimisation of Raman reporter concentration to achieve a stable colloid carried out in Section 4.4.3, antibody conjugation methods were investigated. Covalent antibody conjugation is often considered the most robust approach owing to its chemical permanency, particularly when introduced to biological or pH-variable environments. As such it formed the basis of attempts to conjugate antibodies to p16 in the first instance.

Antibodies are commonly conjugated via their primary amine (-NH_2) groups. Typically the positive charge of the group causes amines to present at the outer surface of the protein structure, theoretically leading to minimal functional disruption through protein denaturing following conjugation. Multiple amine-reactive groups exist with different functionalities, including NHS-esters, imidoesters and carbodiimides.

Taking advantage of the different reactive species available, molecular spacers can be introduced to the covalent conjugation scheme. This reduces the effects of stearic hindrance on antibody capture efficiency when immobilised in large numbers on the surface of a nanoparticle and can also double as a surface blocking agent, although it also increases conjugate size (Figure 4.17).

In this context, a functionalised polyethylene glycol (PEG) molecule acted as a spacer between the nanoparticle surface and the binding ligand. It consists of a 12-unit PEG chain (47.8Å) that is thiol- and carboxyl-terminated (Figure 4.17). The carboxylic acid group is covalently bound to amine groups on the antibody via a carbodiimide activator, N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC). Water-soluble N-hydroxysuccinimide (sulfo-NHS) is added to this reaction scheme to create a more stable intermediate (Figure 4.18).

4.6.2 Comparison of EDC/NHS methods

Conjugation of primary amines using EDC/NHS chemistry can take two forms: 'on-NP' conjugation in which the spacer, conjugation moieties and binding ligand are added layer by layer to the NP surface; and 'off-NP' conjugation where the spacer and ligand are assembled separately before binding to the NP via the thiol group.

On-NP conjugation samples were produced using the method detailed in Section 4.3.3. A commercial monoclonal antibody to human p16 (Santa Cruz, see Section 4.5.1 for details) was used with BPE-coated NPs (1.25µM initial concentration) and

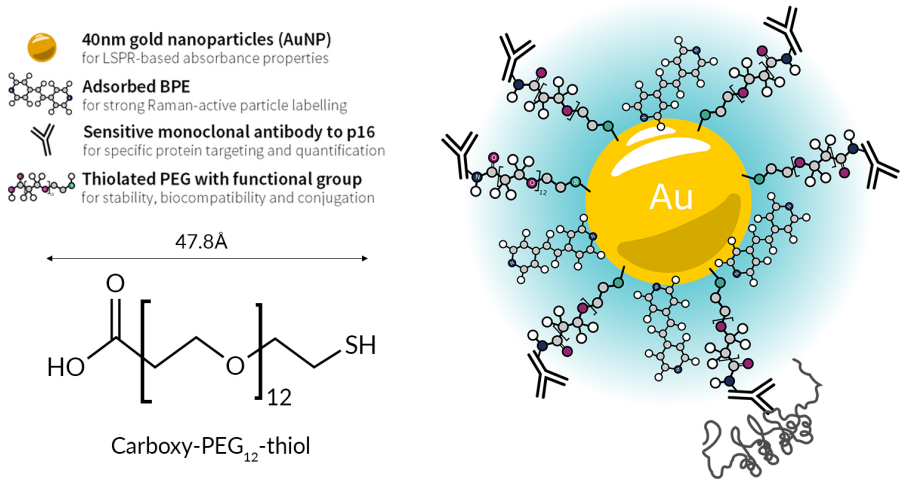


FIGURE 4.17: Schematic of the covalent conjugation system, in which antibodies are attached by their amine groups to the gold spheres via a 12-mol PEG spacer with carboxylated and thiolated end groups.

another IgG1 antibody specific for MDM2 (in-house product, '4B2') used with PYOT-NPs (2.50µM initial concentration). In both cases, the full conjugate was compared to a reporter-only NP solution and a PEG-only control, which followed the full conjugation procedure using distilled water instead of antibody (Figure 4.19). The 4B2 antibody acted as a non-p16 specific isotype control until a genuine IgG1 control was acquired. The 'isotype' sample demonstrated a similar hydrodynamic diameter to the PEG-only control as well as evidence of successful conjugation through a red-shift in the absorbance spectrum compared to bare gold. However, the p16-targeted sample partially aggregated at the same concentration, indicating that the optimised covalent method cannot be applied universally to antibodies of the same isotype. This aggregation took the form of small but visible strands of particles in a near-transparent solution.

The on-NP method was then trialled again with the Santa Cruz monoclonal antibody, introducing variation of antibody concentration. The absorbance measurements and physicochemical characteristics can be seen in Figure 4.20 and 4.21 respectively. Again, all three antibody-coated samples aggregated, with the highest concentration of antibody producing strands of particles.

At this stage, the off-NP method was investigated on the basis that the strand-like aggregation could indicate optimal reagent concentrations with excessive centrifugation. The resultant clumping effect could be mitigated by the use of a method with

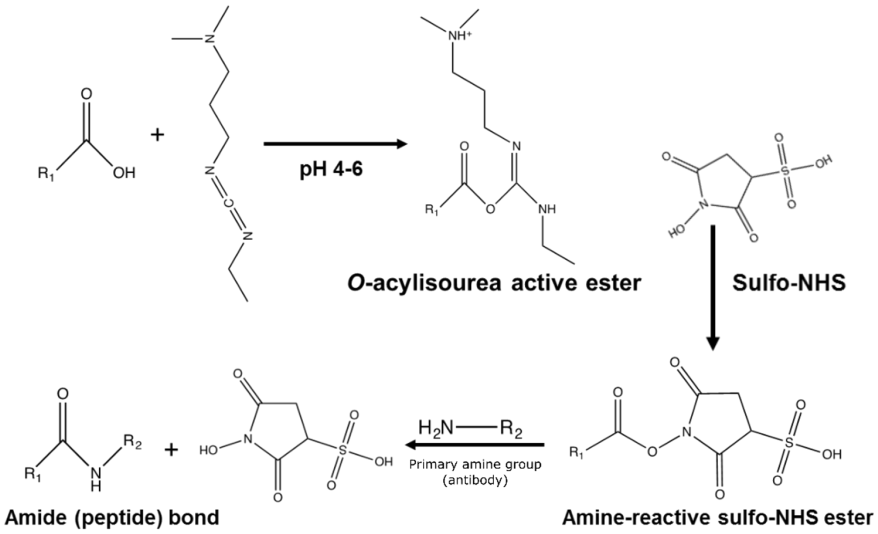


FIGURE 4.18: The reaction scheme for EDC/sulfo-NHS coupling of primary amine groups (antibody) to carboxyl groups (PEG chain). EDC produces an active ester with the carboxyl group which is stabilised by the addition of sulfo-NHS. The introduction of an amine group results in a stable amide bond and sulfo-NHS by-product which is removed with the supernatant after centrifugation.

fewer centrifugation steps (see Figure 4.6). At this stage, a new monoclonal antibody to p16 was released by BD Pharmingen and assessed for efficacy in Section 4.5.1. For continuity both this and the Santa Cruz antibody, now discontinued, were assessed for conjugation stability using the concentrations given in the original protocol (Figure 4.22). The zeta potential measurements in particular suggest that the off-NP method is indeed more stable. Having established that this method was less prone to inducing particle aggregation, a comparison of the antibody types was made and the new ligand found to be more stable. In addition, its increase with size compared to the PEG-only control indicates antibody coverage.

Lower concentration studies with the p16-specific antibody indicated that a reduced antibody concentration results in greater stability. A direct comparison of the lowest concentration was made using the on-NP method, which aggregated into visible strands as before despite demonstrating a higher absolute zeta potential than the off-NP method. Repeats of the off-NP approach with both the BD Pharmingen mAb and genuine isotype control also partially aggregated. At this stage, passive conjugation was adopted as a technique to ascertain the order of magnitude of antibody concentration which could be applied stably to the nanoparticles without consideration of

covalent conjugation kinetics. Optimisation of the passive conjugation technique and accompanying concentration measurements are detailed in Section 4.7. These results indicated that the IgG1 molecules in question should be conjugated in concentrations over an order of magnitude lower than that suggested originally.

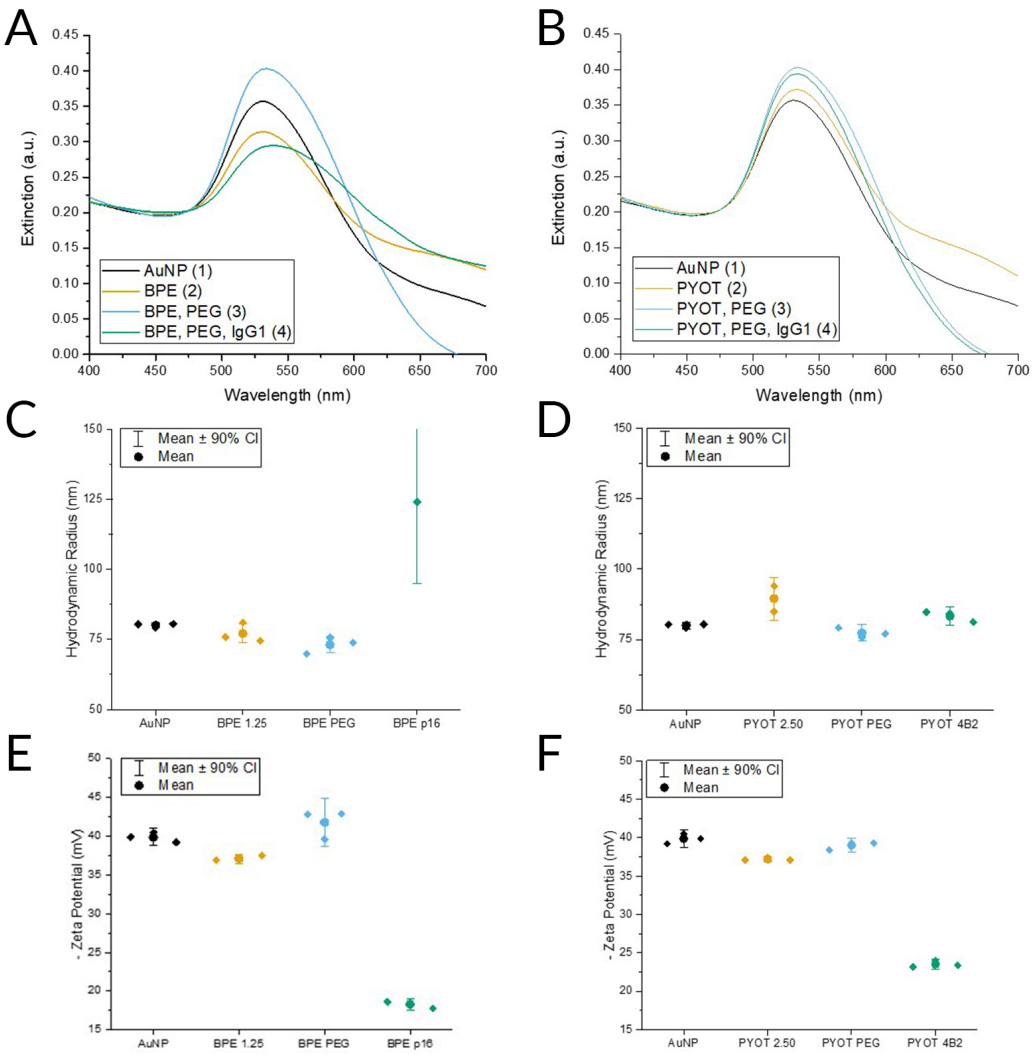


FIGURE 4.19: Targeted and dummy covalent conjugate characteristics at each stage of layering using on-NP EDC-NHS chemistry. Absorbance properties of the A) target (p16 mAb) and B) dummy (IgG1) conjugates; hydrodynamic diameters of the C) target and D) dummy conjugates; zeta (surface charge) potentials of the E) target and F) dummy conjugates.

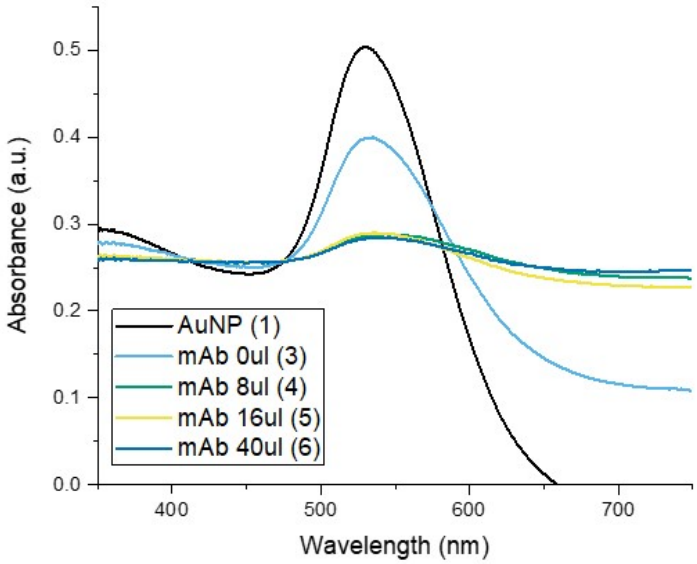


FIGURE 4.20: Absorbance characteristics of on-NP covalent conjugates with Santa Cruz p16 antibody and PEG-only control, showing aggregation of the particles.

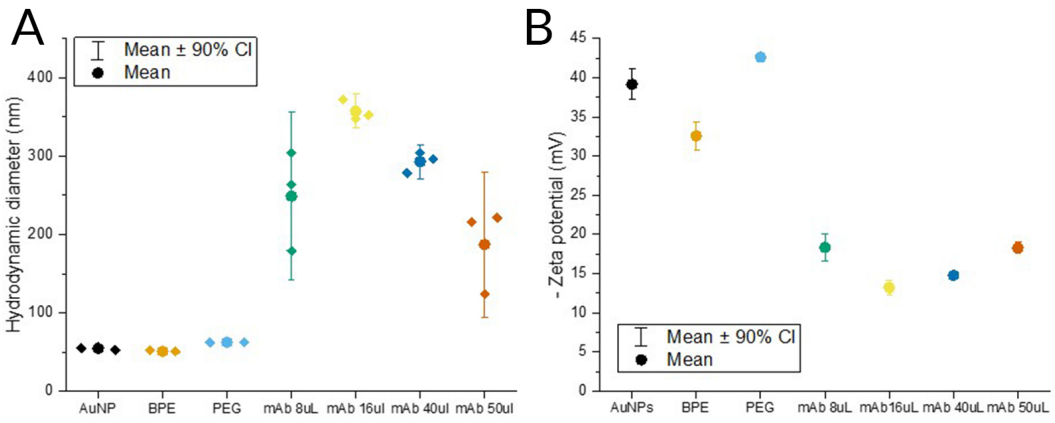


FIGURE 4.21: A) Hydrodynamic diameter and B) zeta potential of the covalent conjugates shown in Figure 4.20, which indicate a lower stability of particle and subsequently larger particle aggregation for the lowest and highest concentrations of antibody used.

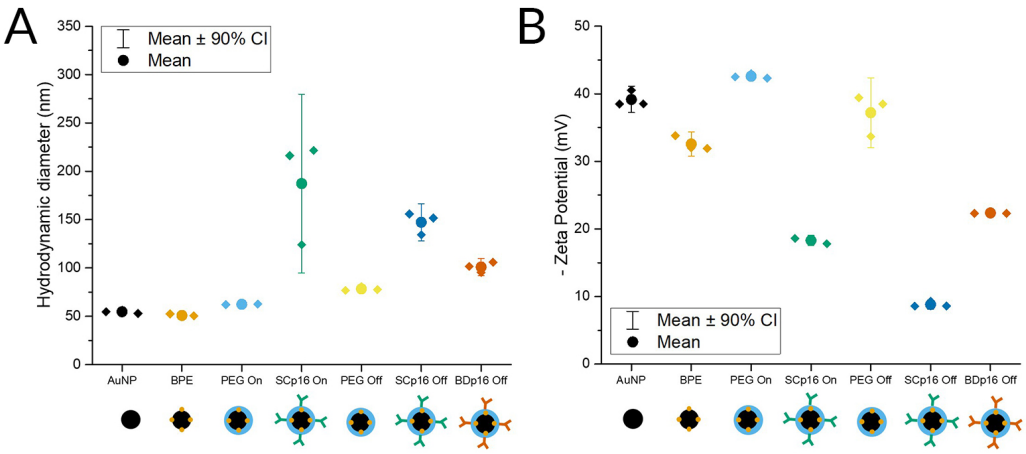


FIGURE 4.22: A) Hydrodynamic diameter and B) zeta potential of covalent conjugates produced using both methods and the Santa Cruz and BD Pharmingen antibodies using the original protocol as stated in Section 4.3.3. Both sets of measurements indicate that the Santa Cruz antibody produces a less stable conjugate than the new BD Pharmingen ligand.

4.6.3 Optimisation of reagent concentration

As a result, twelve conditions were produced to investigate the effects of individual reagent variation on covalently conjugated particles. As seen in Table 4.4, the conditions were chosen in order to explore the effects of antibody concentration variation, equimolar variation of all reagents, coupling reagent variation without antibody, and coupling architecture with a constant carboxyl-terminated PEG layer. These were compared with bare nanoparticles, BPE-NPs and a PEG control. All Raman-active samples used an initial concentration of 1.25µM BPE, equivalent to 12.5nM in NP bulk solution. All conjugation reagents were produced in solution immediately prior to conjugation and stored on ice until required to reduce likelihood of hydrolysis. The absorbance data, peak-to-650nm and wavelength shift values of all samples are represented in Figure 4.23.

Conditions 1 and 10, which represented repeats of the original protocol, both aggregated immediately upon combination of the reporter-NPs with the PEG-mAb complex.

In the first set, the concentration of antibody was varied, with PEG coverage and cross-linking compounds kept constant. In the second, an equimolar variation of all reagents was considered, with '1:1' representing the original protocol. In the third set, the cross-linking compounds were varied ratiometrically with antibody, with the PEG layer kept constant (Figure 4.24).

The absorbance data from the antibody-free controls (Figure 4.24 D) indicate that the original concentration of PEG, and indeed EDC/NHS, provides a greater stability and monodispersity of conjugates than lower concentrations. However, the addition of equal concentrations of monoclonal p16 antibody to that used in the original protocol induced aggregation. This suggests that the stability of the particles is strongly affected by the specific antibody to be conjugated, not merely its type. Out of the conjugation reagents, concentration of PEG appears to have the greatest impact on LSPR wavelength shift (Samples 7, 8, 9), which is amplified by the addition of antibodies (Samples 1, 5, 6).

These same samples were then introduced to commercial dipstick assays (see Section 2.6.2), used to assess the correct conjugation of antibody onto the surface. The assays are printed with a Protein A/G mix which binds to the Fc portion of the antibody, causing collection of any successfully conjugated nanoparticles onto the capture strip which can be seen by eye due to the absorbance properties of the NPs themselves. The variation of antibody alone results in a high background with no signal at the capture line, indicating that the non-conjugated activated PEG is interacting with

Condition	**CT-PEG (μM)**	**EDC/NHS (μg/ml)**	**mAb (μg/ml)**	**HEPES (mM)**
AuNP	-	-	-	-
BPE	-	-	-	-
1	1000	1000	500	10
2	1000	1000	5	10
3	1000	1000	0.5	10
4	1000	-	-	10
5	10	10	5	1
6	1	1	0.5	0.1
7	1000	1000	-	10
8	10	10	-	1
9	1	1	-	0.1
10(1)	1000	1000	500	10
11	1000	10	5	10
12	1000	1	0.5	10

TABLE 4.4: The concentrations of CT-PEG (blocking polymer and functionalised linker), EDC/NHS (functionalisation reagents), antibody (binging ligand) and HEPES (conjugation buffer) used for an experiment investigating covalent conjugation stability with the antibody in question.

the nitrocellulose membrane. The pure AuNP and PEG-constant particles do not produce any colorimetric signal, showing that a positive lateral flow test is the result of successful antibody conjugation and not activated carboxyl groups or non-specific interaction of the Protein A/G mix with a bare metal surface. The ratiometrically varied samples produce a relatively clean background and signal at the capture line (Figure 4.25).

None of the above samples produced a visible colorimetric signal with a purified protein dot blot, indicating that the technique used successfully conjugates the antibody but disrupts its functionality. Multiple attempts to reproduce the successful conjugation conditions above failed, with no signal seen on commercial dipstick assays at the test line.

A set of samples representing a full investigation of equimolar reagent concentration was then produced, using 0.75μM BPE as the initial concentration of Raman reporters. The reagent concentrations and dipstick assay results can be seen in Figure 4.26. Unlike the higher-concentration samples produced previously, this repeat remained stable and produced a strong positive signal at the Protein A/G capture line. This drops off for the tenfold dilution and then increases again for Conditions 5 and 6 as seen previously. This indicates that the positive test line seen previously may

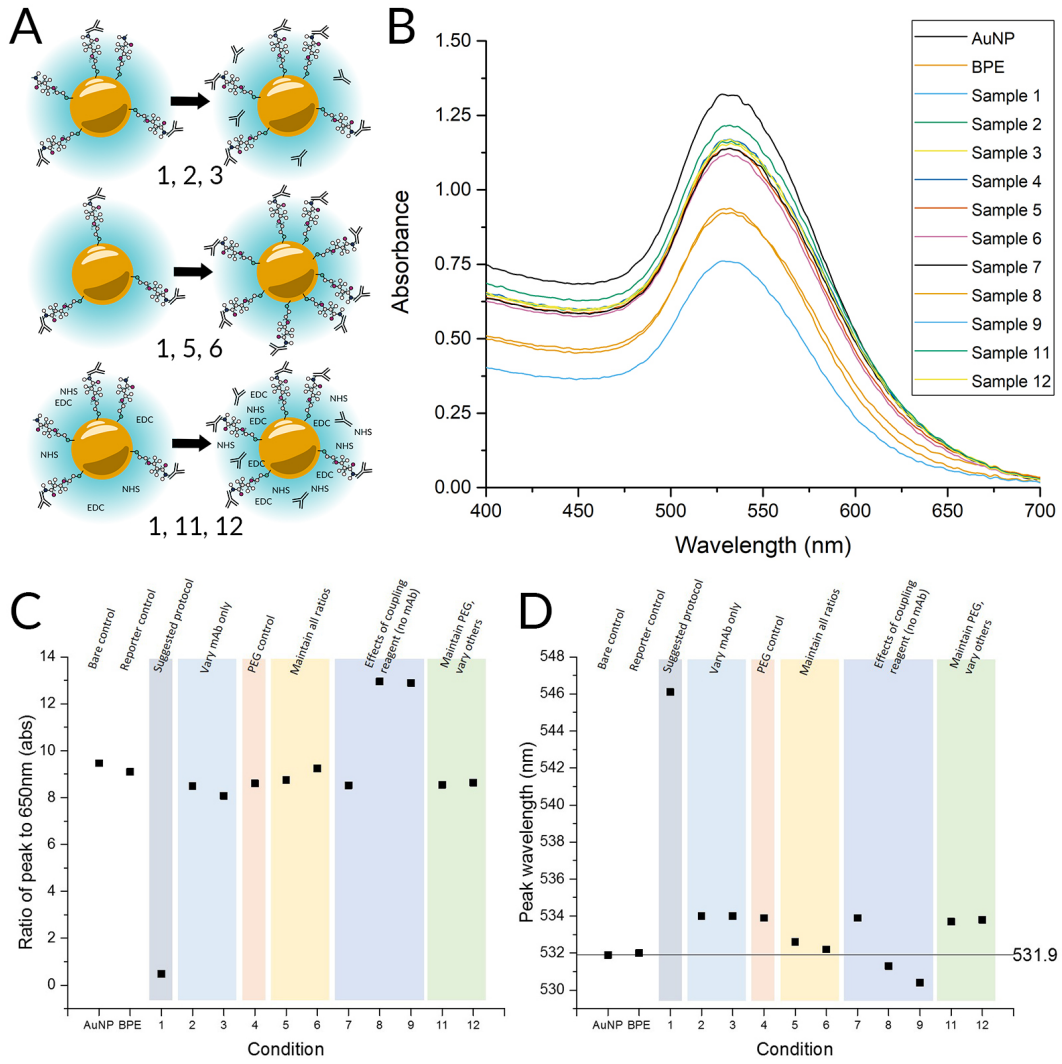


FIGURE 4.23: A) Schematic of the conditions explored for this experiment, including antibody concentration, all reagents, and all reagents except PEG; B) the extinction characteristics of the samples; C) ratio of peak height to extinction at 650nm, providing a concentration-free indication of aggregation status; D) extinction peak wavelength, which shifts fractionally (1-4nm) with successful surface modification but significantly with unwanted aggregation.

be the result of insufficient blocking on the particle surface rather than the number of antibodies successfully conjugated. However, despite the indication of successful conjugation for Condition '3', dot blots with purified p16 did not yield a positive signal as they did for passive and bioconjugated nanoparticle-antibody complexes.

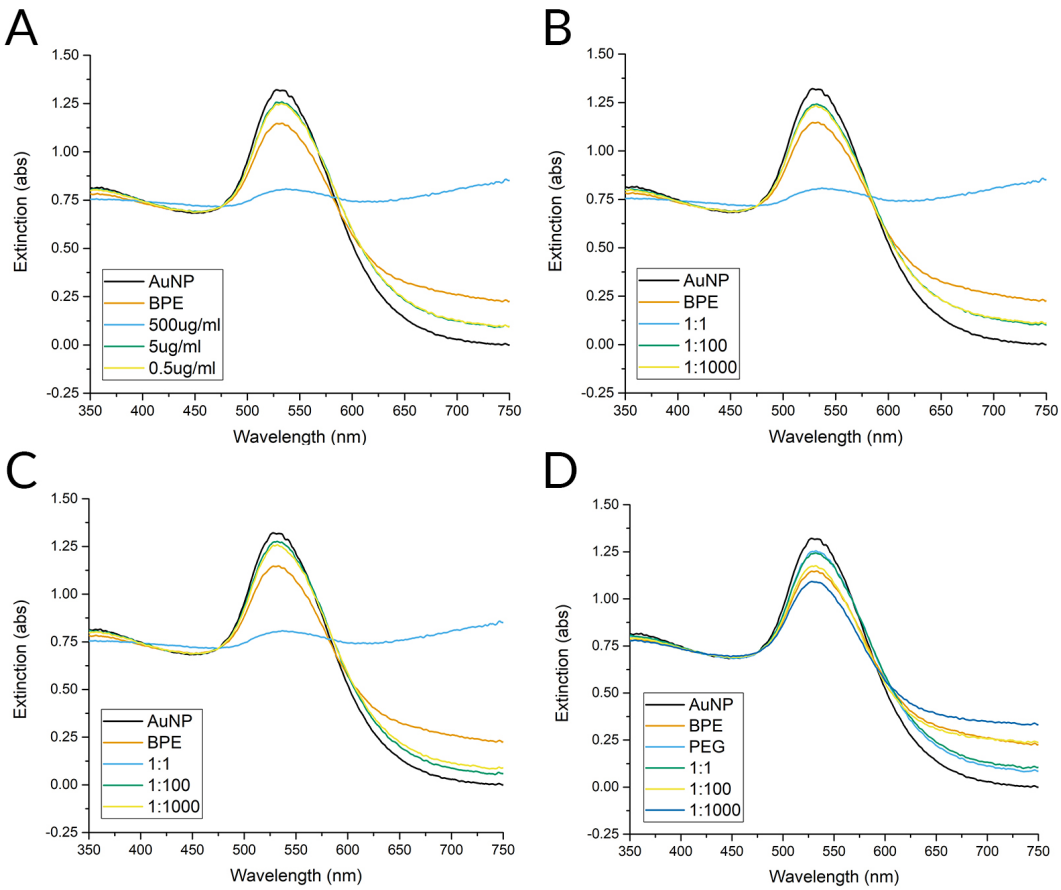


FIGURE 4.24: Extinction spectra of conjugates with different concentrations of A) antibody; B) all reagents; C) all reagents except PEG; D) all reagents with no antibody.

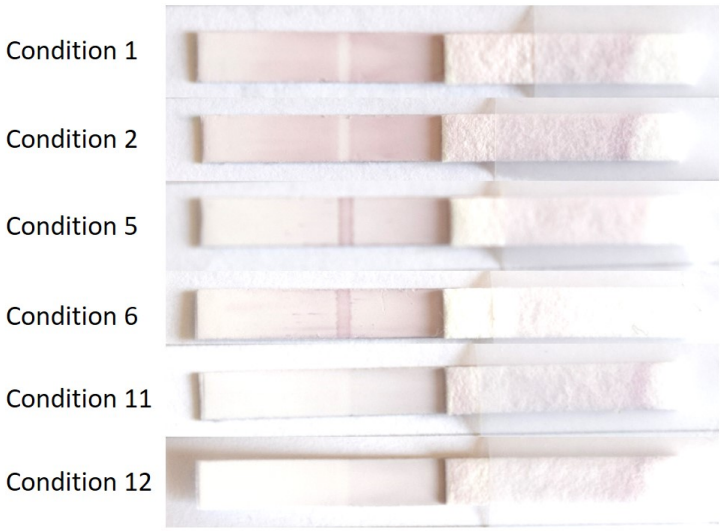


FIGURE 4.25: The commercial dipstick assay results for conditions where antibody was used. Conditions 1 and 2 varied mAb only, 5 and 6 varied all reagents equally, and conditions 11 and 12 varied all reagents except PEG.

Condition	PEG	EDC/NHS	mAb	HEPES	
3 (neat)	1mM	1mg/mL	0.5mg/mL	10mM	
4 (1:10)	100uM	100ug/mL	50ug/mL	1mM	
5 (1:100)	10uM	10ug/mL	5ug/mL	100uM	
6 (1:1000)	1uM	1ug/mL	0.5ug/mL	10uM	

FIGURE 4.26: The commercial dipstick assay results for covalent conjugation conditions where the original ratio of reagents (EDC/NHS, CT-PEG and mAb) were retained but serially diluted.

4.6.4 Variability of binding functionality

Following the optimisation of reporter signal seen in Section 4.8.4, a new set of covalent conjugates were produced as a final assessment of their use in Raman-based diagnostic applications. Both on-NP and off-NP methods were trialled, with 0.75µM BPE as in the previous set of samples as well as 5µM BPE which was found to provide stable, highly Raman-active conjugates.

1ml samples of 1:1, 1:10, 1:100 and 1:1000 dilutions of the protocols described in Section 4.3.3 were produced in addition to a 1:1 antibody-free control with both reporter concentrations, where the antibody volume was replaced with distilled water. Prior to the final centrifugation and resuspension in pH7.6 phosphate buffer, the samples were split into two 500µl volumes. Following centrifugation and supernatant removal, one of the samples was resuspended in 480µl phosphate buffer for optical characterisation (Figure 4.27) and the other in 80µl phosphate buffer for interrogation with commercial dipstick assays (Figure 4.28). The assays were prepared with 0.5µl purified p16 protein spots (100ng/µl in PBS) and left to dry prior to assay use.

Both sets of on-NP samples produced at 1:100 and 1:1000 reagent dilution aggregated following resuspension in HEPES and addition of antibody, as can be seen from the heavily reduced absorbance characteristics in Figure 4.27. For these measurements, acquired immediately after production, the off-NP samples demonstrate the same trend but not to the same extent of aggregation. However after two days stored at room temperature, the off-NP 1:100 and 1:1000 samples also aggregated.

The dipstick assays appear to demonstrate stronger target binding characteristics for the off-NP conjugates despite a lower test-line signal. This suggests that the orientation of antibodies in this case is more 'outward-facing' or that this method is less functionally disruptive.

The Raman spectra of these conjugates, acquired after off-NP sample aggregation had taken place, can be seen in Figure 4.29. The off-NP method produced a stronger Raman signal than on-NP for both concentrations of reporter. However, this technique still yields a lower Raman signal than stable bioconjugates and uses much higher quantities of antibody.

Owing to the intrinsically inconsistent nature of this technique, the lower Raman signal which can be achieved using the same concentration of reporter, and the increased cost of reagents which must be considered in the context of population health applications, the covalent conjugation method was not chosen for production of SERS conjugations for cell-based diagnosis.

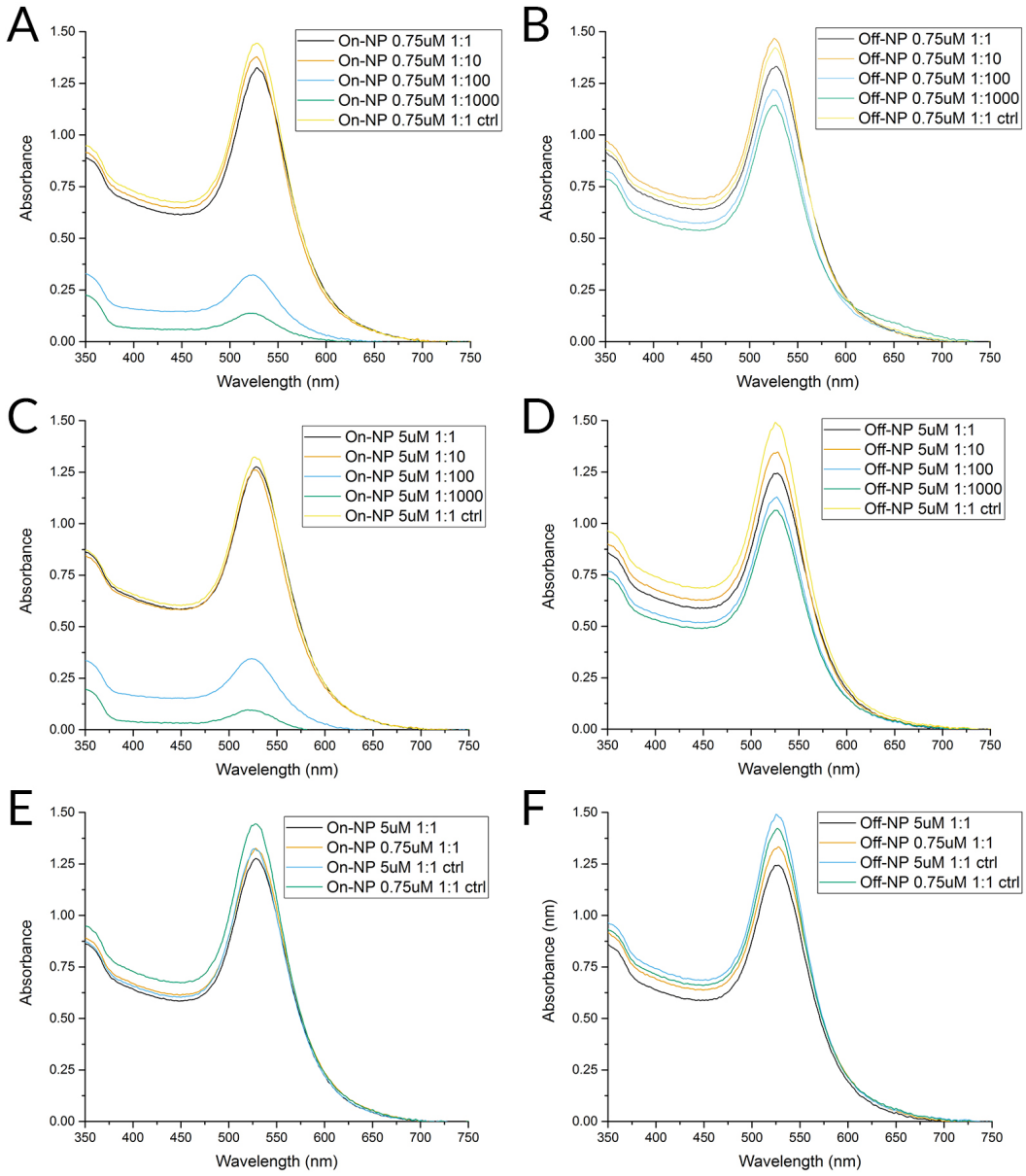


FIGURE 4.27: Extinction spectra for covalent conjugates produced using A) on-NP method with 0.75µM BPE-NPs; B) off-NP method with 0.75µM BPE-NPs; C) on-NP method with 5.00µM BPE-NPs; D) off-NP method with 5.00µM BPE-NPs. Comparison of E) on-NP samples; F) off-NP samples.

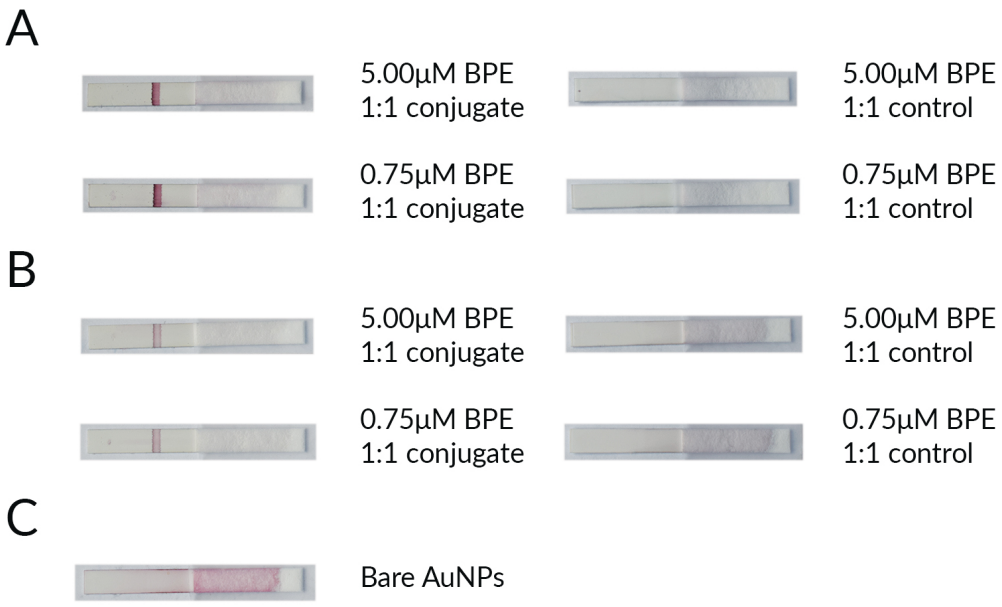


FIGURE 4.28: Dipstick assay binding sensitivity of 5x concentrated EDC/NHS conjugates. A) On-NP method. B) Off-NP method. C) Bare AuNPs. Faint binding spots can be seen for all of the non-control items (left).

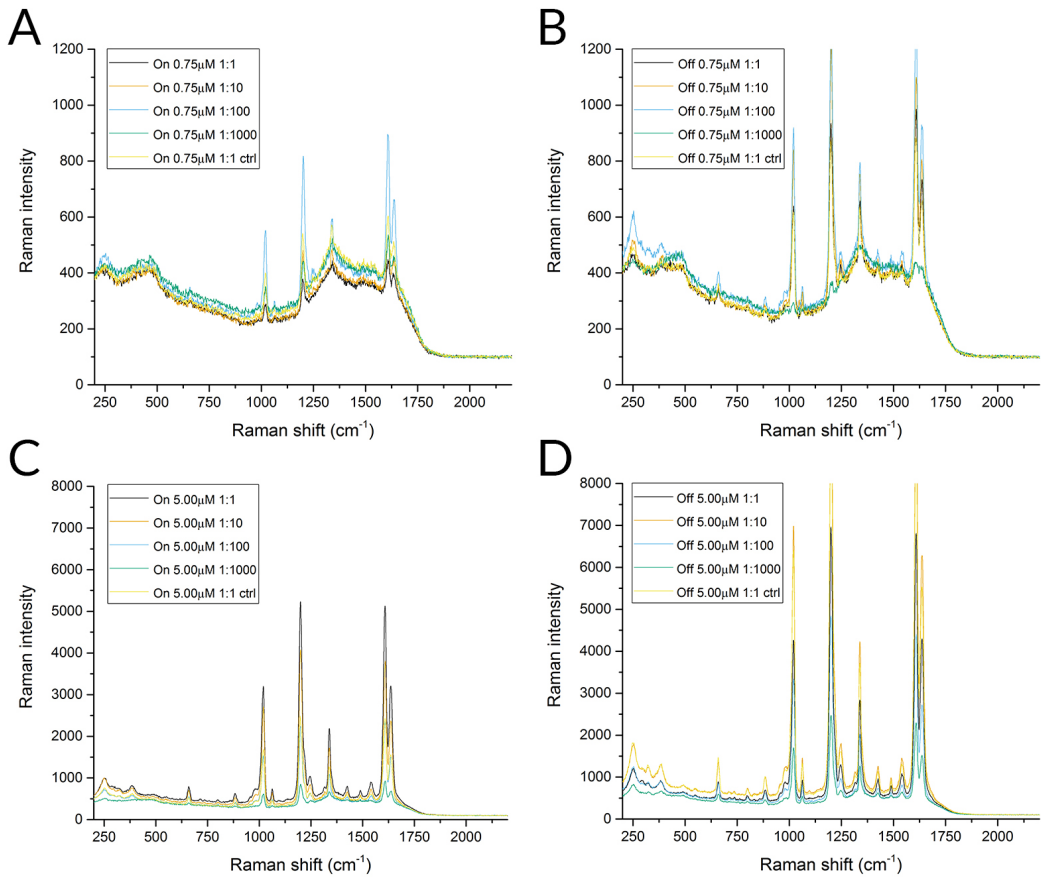


FIGURE 4.29: Raman spectra of the samples produced and shown in Figure 4.27.

4.7 Passive conjugation of antibody

4.7.1 Passive conjugation scheme

Passive antibody conjugation is widely considered to be the simplest of conjugation methods as it does not require additional coupling moieties beyond the binding ligand and nanoparticle surface. Instead the technique relies upon electrostatic, van der Waals and hydrophobic interactions between the two components. The basic premise is that by altering the pH of the nanoparticle buffer to match the isoelectric point (pI) of the antibody so that it carries no net charge, attractive forces will cause the antibody to passively adsorb onto the nanoparticle surface (Figure 4.30). As such, conjugates can be produced rapidly and inexpensively without iterative optimisation of additional components.

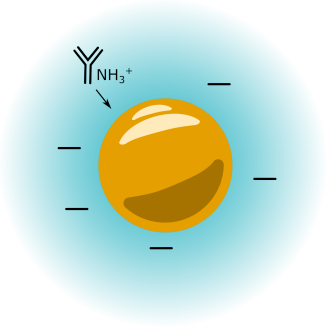


FIGURE 4.30: Schematic of the passive conjugation mechanism, in which the pH of the surrounding medium is shifted so that proteins may overcome the negative surface potential of the gold spheres and adsorb onto the surface.

Despite its theoretical advantages the reality of this interaction can be complex. Both nanoparticles and antibodies are sensitive to changes in pH. It is difficult to ascertain the exact pI of a commercial antibody with a protected amino acid sequence. Successful passive conjugation can lead to disruption of the tertiary protein structure, leading to a loss of biological functionality. The aggregation of nanoparticles is also dependent on the relative coverage of Raman reporters and binding ligands - insufficient coverage and the bare NP surface will be subject to the aggregating effects of buffer ions, but too high a concentration and they may clump together. The use of blocking agents such as bovine serum albumin (BSA) to passivate the bare surface must not interfere with the active sites of the ligand or interact non-specifically with the environment. Reporters and antibodies compete for space on the surface and must

be finely tuned to ensure that the final conjugate is both optically and biologically active.

4.7.2 Optimisation of antibody coverage and surface competition

Passive conjugation was initially explored as both a potential alternative to covalent conjugation which may circumvent issues over binding non-functionality due to reagent cross-linking, and as a way of rapidly ascertaining the optimal concentration of antibody to use with 40nm AuNPs. The pH of the Raman active reporters used in Sections 4.4.2 and 4.6.2 for 1.25µM BPE and 2.50µM PYOT was 8.5 and 7.9 respectively, measured using a digital pH probe with three-point calibration (Eutech Instruments). The isoelectric point of IgG1 antibodies is typically between 8.2 and 9.0. If the surrounding environment has a lower pH than this, then the antibody will carry a net positive charge. If higher then it will carry a net negative charge, causing electrostatic repulsion between the ligand and negatively charged gold colloid. Passively conjugated 'target' (p16 mAb, BPE) and 'dummy' (IgG1 isotype, PYOT) particles were produced using the initial protocol detailed in Section 4.3.4 and found to be stable. The conjugates were then applied to a dipstick assay as outlined in Section 2.6.2, prepared with 0.5µl purified p16 protein. The assays appeared to indicate a positive line for the target conjugates but not at the original application site of the protein, and so the result was deemed inconclusive. Additional samples produced to the same protocol and tested using a dot blot method did not bind, and indeed appeared to aggregate into the nitrocellulose membrane around the site of the protein spot, indicating a problem with flow and blocking.

It was then suggested that I should investigate the application of antibodies to the surface prior to the Raman reporter, as colleagues had seen greater success with this approach. These were produced according to the following protocol:

1. Centrifuge 500µl AuNPs at 2300g for 15 minutes, remove 480µl supernatant and resuspend in equal volume of pH7.9 phosphate buffer.

2. Add 5µl antibody or volume as desired and leave to shake for 15 minutes at 500rpm.

3. Add 5µl 10%[w/v] BSA in PBS and shake for 30 minutes.

4. Centrifuge at 2300g for 15 minutes, remove 480µl supernatant and resuspend in equal volume of pH7.0 phosphate buffer.

5. Add 5µl appropriately diluted Raman reporter solution and shake for 30 minutes.

6. Centrifuge at 2300g for 15 minutes, remove 480µl supernatant and resuspend in equal volume of pH7.0 phosphate buffer.

These samples aggregated. At this stage sample production was focused down to dummy particle production only in order to conserve reagents. Comparisons were made with bare AuNPs and PYOT-NPs passively conjugated with IgG1 isotype control antibody (1.0mg/ml) using 1, 2, 3, 4 and 5µl antibody solution with 500µl particle sample, equivalent to 2-10µg/ml in 2µg/ml intervals. Of the bare AuNP particles, the addition of 1,2 and 3µl IgG1 resulted in complete aggregation, which indicated insufficient coverage. Addition of 4 and 5µl caused only partial aggregation, whereas all of the conjugated PYOT-NP samples remained stable. All samples were stored at 4°C overnight, causing complete aggregation of all conjugates except the AuNP-only 5µl sample. Further samples up to 26µg/ml were produced, characterised optically as shown in Figure 4.31. No trend could be found for this data - concentrations 6-20µg/ml gave the greatest particle stability but where not found to bind to p16.

4.7.3 Effect of pH on passive conjugates

At this stage it was discovered that bare AuNPs, with an intrinsic pH4.0, aggregate overnight when pelleted and resuspended in pH8.0 phosphate buffer as though for conjugation. This was retested four times and found to be a consistent effect. As a result of this, all further passive conjugation experiments were carried out using manually pH-adjusted particles. Buffer exchange using dialysis tubing (Spectra/Por, 12-14kDa) was also found to aggregate the particles. The manually adjusted AuNPs (now pH8.5 through dropwise addition of 0.5M NaOH) were used with ten-fold dilutions of both p16 mAb and IgG1 isotype control antibody to give characteristics as shown in Figure 4.32. For both p16 mAb and its IgG1 isotype, the peak extinction wavelength indicates an excess of antibody from 1000ng/ml and above. The aggregation shift across the range tested is more pronounced for IgG1 isotype control, with an absorbance drop and wavelength redshift on the order of 0.5 and 10nm respectively, compared to 0.1 and 5nm for p16 mAb.

Additional experiments using 10ng/ml final concentration of either p16 mAb or IgG1 control were produced using AuNPs adjusted to pH7.1, 8.0, 9.1 and 10.0, along with BSA-only controls. All were stable but none found to give a positive test-line signal with the commercial dipstick assay 'Conjugate Check & Go!'. Dot blot assays

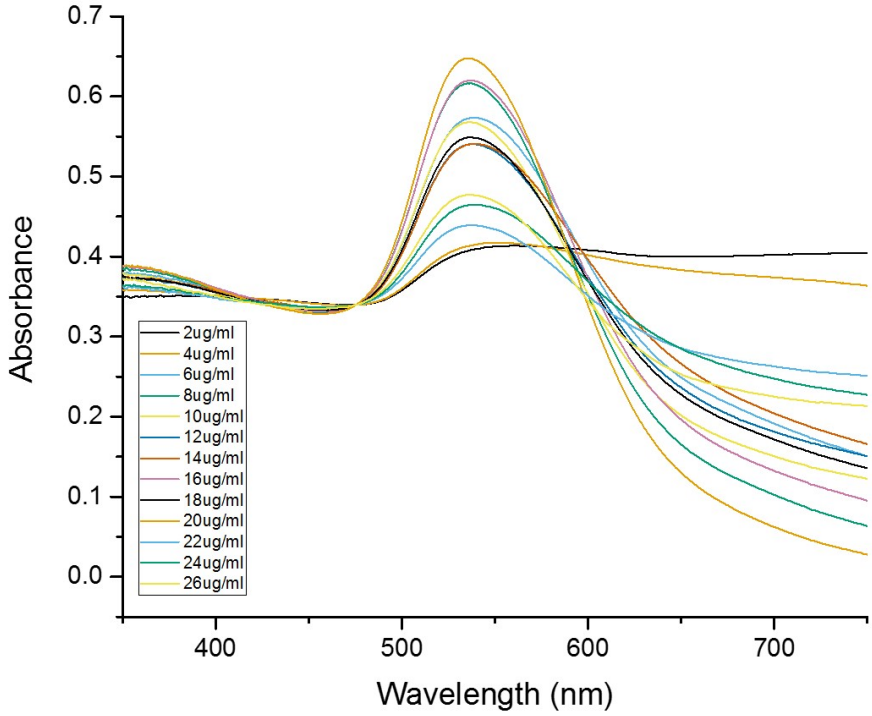


FIGURE 4.31: Extinction spectra for different concentrations of antibody (final concentration) passively conjugated onto 40nm gold nanoparticles.

with the same samples also failed. Bulk Raman measurements also indicated that no Raman reporters were adsorbing onto the surface following passive conjugation of the binding ligands. As such, a set of samples produced according to the above protocol (AuNPs pH8.5) were compared to samples produced with simultaneous ligand and reporter addition at both pH4.5 and pH8.5. All were stable but none found to produce a signal with an in-house dipstick assay.

It was then calculated that for the stable passive conjugates (binding ligand at 100ng/ml or below, equivalent to 1:100 dilution ratio), conjugates would have to be concentrated by at least five times to achieve a detectable bulk signal if compared to the antibody-only ELISA performed in Section 4.5.2. This would explain the lack of signal from the conjugates in dipstick assays, despite the optical characteristics indicating successful adsorption of the antibody.

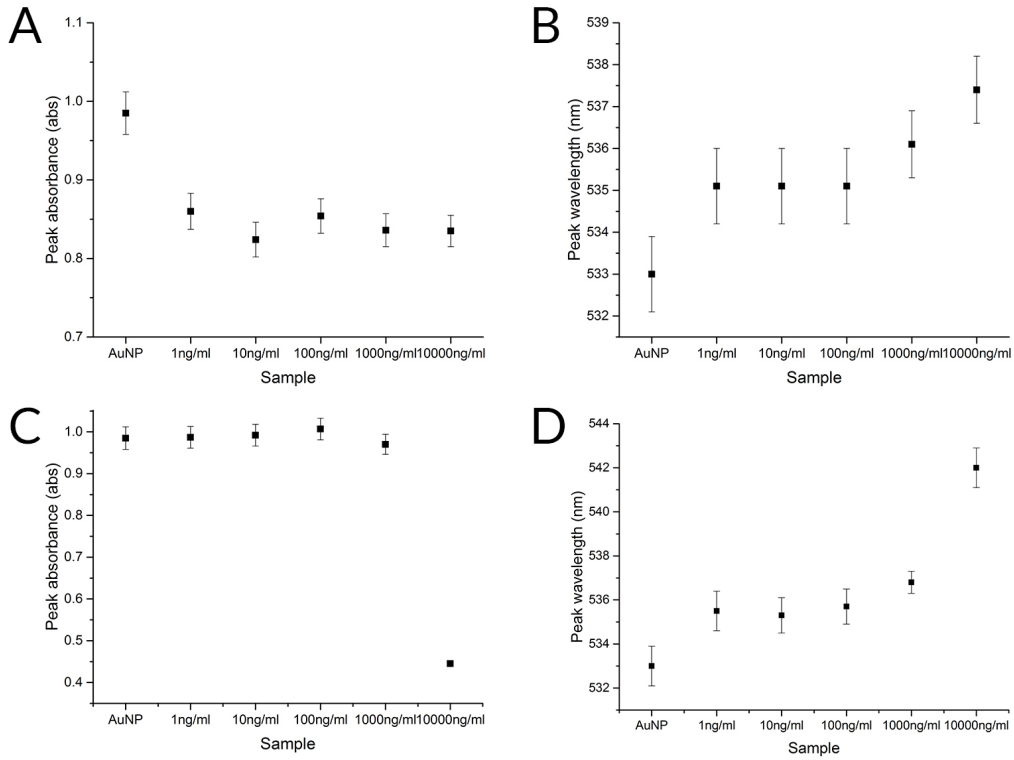


FIGURE 4.32: A) Extinction peak height and B) wavelength for p16 mAb passively conjugated AuNPs. C) Extinction peak height and D) wavelength for IgG1 isotype passively conjugated AuNPs.

4.7.4 Passive conjugates with reduced reporter coverage

In order to maintain conjugate stability it was then decided to reduce the concentration of reporter used for the surface, as this would mitigate surface competition between the two components. Reporter-coated AuNPs using initial concentrations of 0.50, 0.75 and 1.00µM BPE were produced in triplicate according to the following protocol:

1. Dilute 1mM stock (BPE in ethanol, 4°C) in dH_2O appropriately to give desired initial reporter concentrations.
2. Add 5µl reporter solution to 495µl AuNPs and shake at 500rpm for 30 minutes.
3. Centrifuge at 2300g for 15 minutes, remove 480µl supernatant and resuspend in equal volume of pH7.9 phosphate buffer.
4. Store at room temperature for 48 hours to account for mid-term stability and characterise optically.

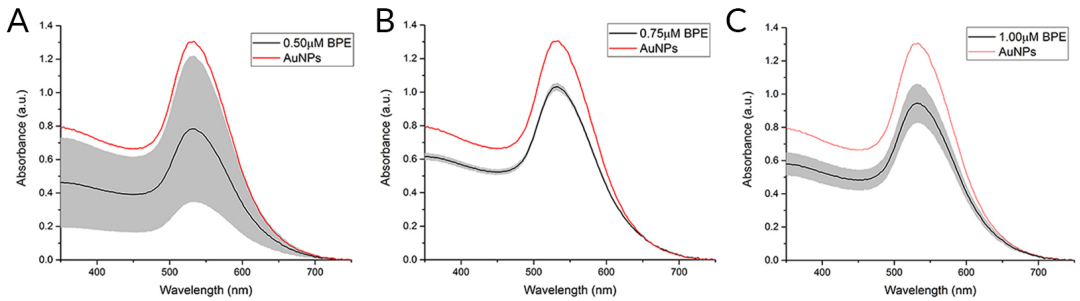


FIGURE 4.33: Extinction spectra in triplicate with an AuNP reference for initial BPE concentrations of A) 0.50µM; B) 0.75µM; C) 1.00µM.

The absorbance profiles of the samples prior to antibody addition can be seen in Figure 4.33, which indicate that 0.75µM BPE coverage provides the most consistent result. The final conjugates also investigated antibody dilution in addition to reporter coverage, and were introduced to commercial dipstick assays to test for successful passive adsorption of antibody. Antibody dilutions of 1:10, 1:100 and 1:1000 were used. The 0.5µM BPE, 1:10 mAb (equivalent to 1µg/ml) sample aggregated, indicating that adding an excess of binding ligand to particles without monolayer reporter coverage can disrupt their stability. The remaining samples, shown in Figure 4.34, demonstrate to an extent the relationship between reporter and binding ligand coverage. Assessments of the greyscale intensity of the red channel of paper assays at the binding spot and bare strip were made using ImageJ.

As expected, the signal strength produced at the test line increases with antibody concentration, so that with 1:10 dilution a higher proportion of particles have antibodies adsorbed onto their surface. As with experiments in Section 4.6.3, 1:1000 mAb dilution results in an incredibly weak test-line signal. When compared to other 1:100 samples, the 1.00µM BPE, 1:100 mAb conjugate also seems to indicate that a higher coverage of reporter requires higher antibody saturation in order to generate a signal. This may be due to the fact that higher coverage of reporter leads to weaker attractive forces between the particle surface and protein, meaning that larger quantities of protein are required to increase likelihood of contact between the two components.

The two optimal samples in terms of colorimetric contrast - 0.75µM and 1.00µM BPE with 1:10 mAb dilution - were produced in triplicate and found to be consistent (Figure 4.35). The 0.75µM BPE sample was also then used with a Western Blot membrane (see Sections 2.3 and 2.7) to test whether it would bind visibly to native p16 from SiHa cells. When compared to HCT116, a colorectal carcinoma cell line regularly used as a negative control for the protein, it was found that the conjugates appear to bind

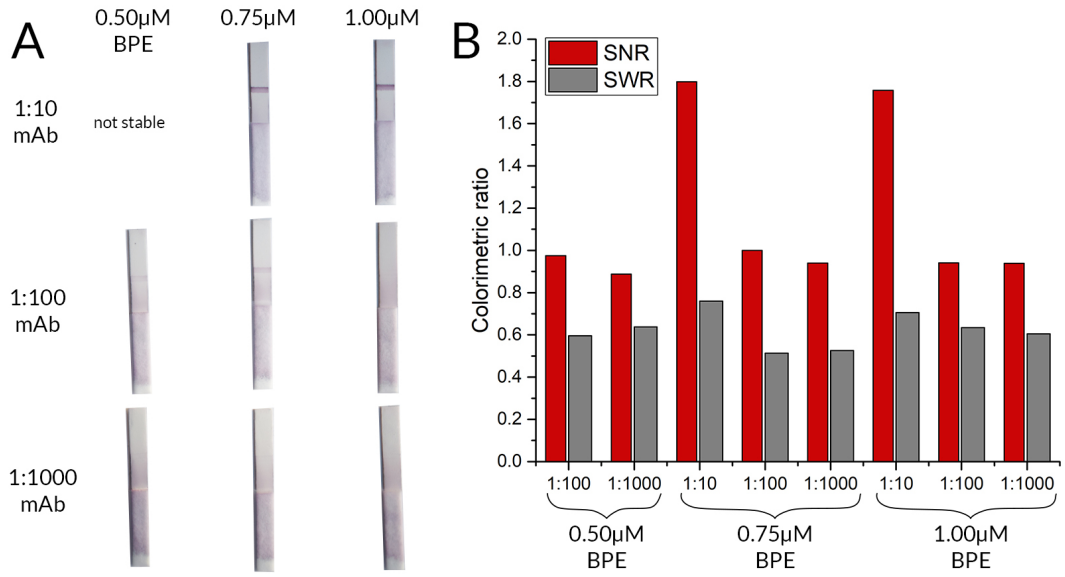


FIGURE 4.34: A) Dipstick assays demonstrating antibody coverage of passive conjugates at different concentrations of BPE and antibody; B) colorimetric assessment of the assays showing test line intensity (SNR) and background accumulation compared to bare strip (SWR).

non-specifically to all proteins on the membrane, possibly as the result of insufficient blocking of the particle surface (also Figure 4.35).

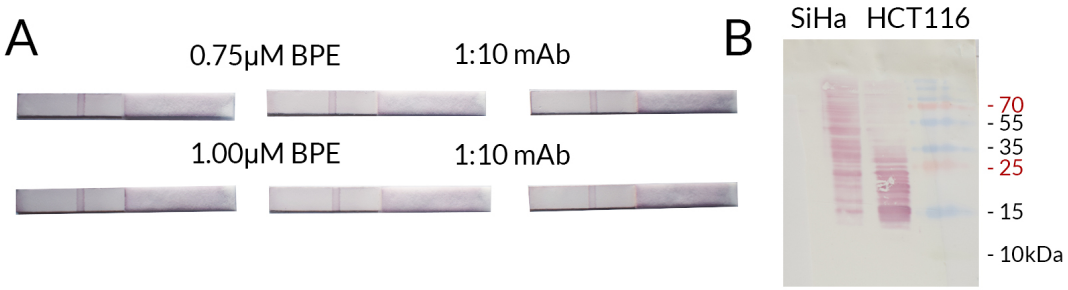


FIGURE 4.35: A) The two best conditions from Figure 4.34 produced in triplicate; B) non-specific binding to membrane-bound proteins for p16-positive (SiHa) and negative (HCT116) cells.

4.7.5 Consideration of passive conjugate blocking on protein interaction

Conjugate samples were then produced according to the following procedure:

1. Add 5µl desired reporter concentration to 495µl AuNPs and shake for 30 minutes at 500rpm.
2. Centrifuge at 2300g for 15 minutes to pellet. Remove 480µl supernatant and suspend in equal volume of 10mM pH7.9 phosphate buffer.
3. Add 5µl desired antibody concentration (1:10 dilution in PBS) to reporter AuNPs and shake for 15 minutes.
4. Add 10µl desired concentration [% w/v] BSA in PBS and shake for 30 minutes.
5. Centrifuge at 2300g for 15 minutes to pellet. Remove supernatant. Resuspend in either 480µl 10mM pH7.6 phosphate buffer for normal concentration or 80µl for a 5x optical density conjugate suitable for blotting.

The BSA content of the blocking solution was varied from 10 to 30% (equivalent to solution saturation) in 5% intervals. The resulting conjugates, along with BPE-coated AuNPs in phosphate buffer, were applied to a Western blot membrane. The non-specific control did not bind to the protein content of the membrane, whereas all five blocked samples did (Figure 4.36). The p16 band was clear in all cases, along with all remaining proteins. There does not appear to be any trend in binding strength, either for p16 or other proteins, related to the percentage weight of BSA in solution.

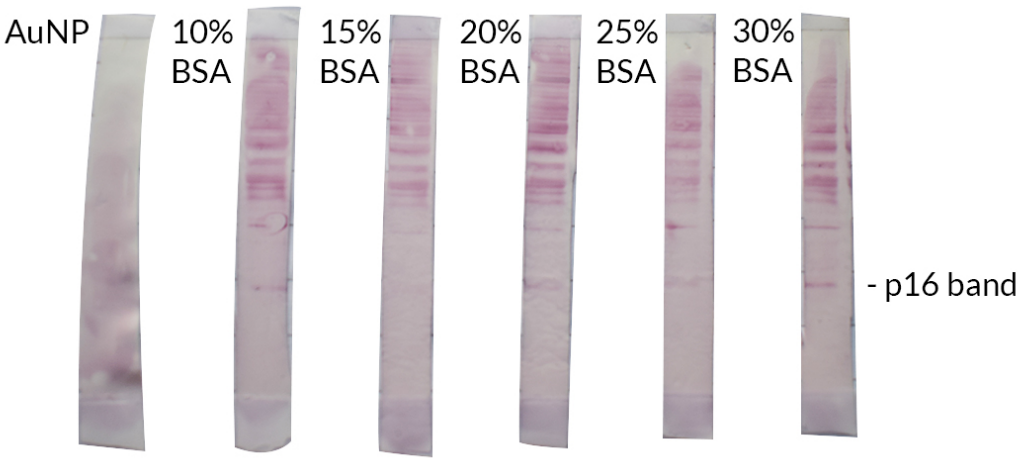


FIGURE 4.36: Different concentrations of BSA used to block passively conjugated nanoparticles, from the standard concentration to solution saturation. Additional block did not aggregate the particles but did not reduce non-specific binding.

A conventional Western Blot (Section 2.3.2) was then run using wild-type SiHa lysate, using 3% BSA in PBS-T (w/v) as a blocking agent instead of 5% fatless milk solution. This resulted in a double-band for p16 in addition to an extremely high background (Figure 4.37). A second Western blot directly comparing the two blocking agents with both antibody and secondary-only control again suggested that BSA is an inappropriate blocking agent to use in the context of blotting methods, due to the high background. Two final Western blots were then produced. The first used BSA as a protein sample and showed that there is no specific reaction between BSA and the antibody used for this project (Figure 4.38). The second, in which passive conjugates were blotted onto a milk-blocked membrane, showed that milk blocking significantly reduces the non-specific background of the conjugates but also reduces the colorimetric signal seen at the p16 band so that it is no longer detectable by eye (also Figure 4.38).

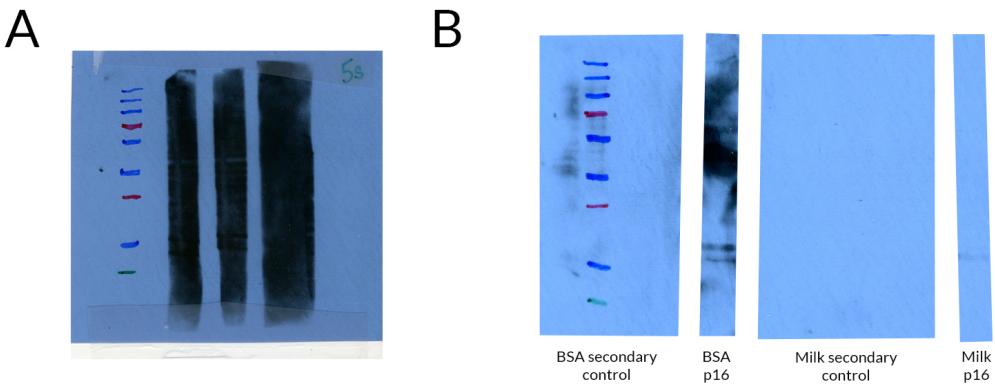


FIGURE 4.37: Western blots using A) BSA as a membrane block (in triplcate), B) BSA and milk as membrane blocking agents, compared with secondary only control. Stringent and identical washing procedures were used for each membrane, indicating a reaction between BSA and the primary antibody.

Casein (VWR International), polyethylene glycol (average m.w. 8000, Sigma Aldrich) and goat serum (Sigma Aldrich) were all investigated as alternative blocking agents. The casein samples aggregated, and remaining samples produced a false positive signal on commercial dipstick assays owing to the need for BSA in the running buffer solution for the assay (Figure 4.39).

Finally, dot blots were produced for the best conditions of all conjugate types to test for binding strength in a colorimetric context. In each case 1µl spots of purified

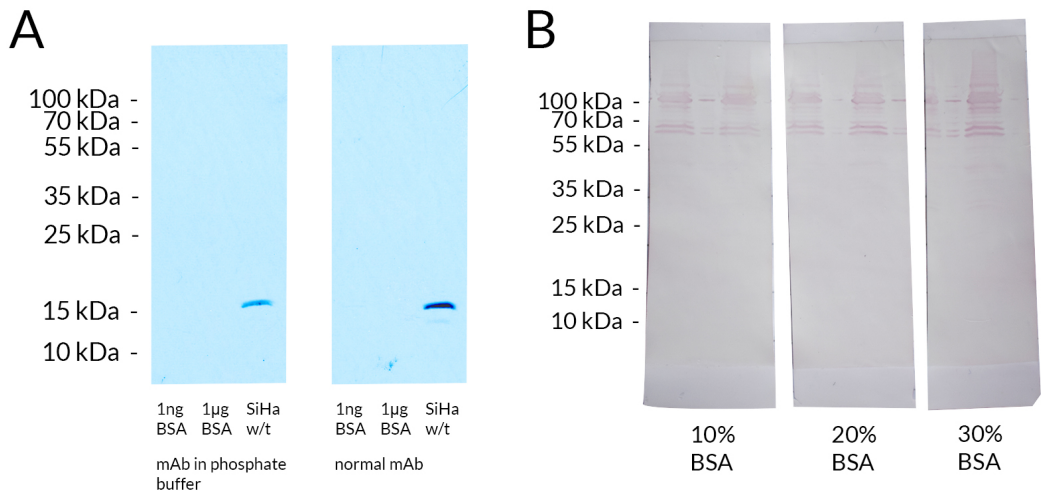


FIGURE 4.38: Western blot membranes showing that A) the reaction with BSA is non-specific and that antibody sensitivity is reduced by storage in phosphate buffer; B) membrane blocking with milk solution gives a marked improvement in non-specific background but does not remove it completely.

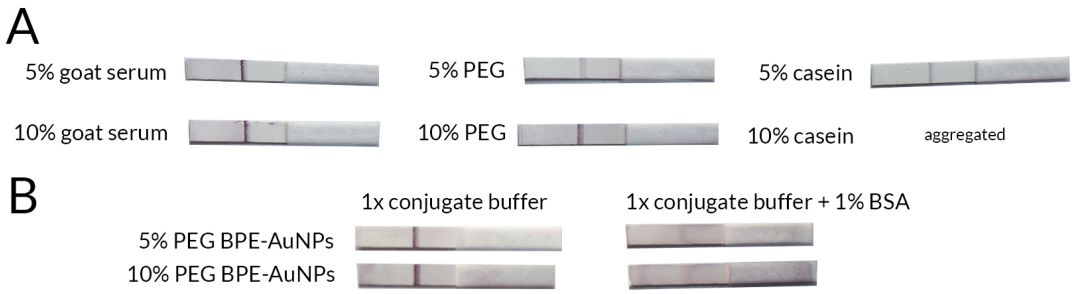


FIGURE 4.39: A) brief investigation of alternative blocking agents with dipstick assays; B) demonstration of requirement for BSA in conjugate buffer as an assay-blocking measure.

p16 protein (100ng/µl and 10ng/µl) were applied to a PBS-wetted nitrocellulose membrane and left to dry, before the membrane was blocked for 30 minutes at room temperature using 1% [w/v] fatless milk powder in PBS. Conjugates were incubated for two hours before being washed with milk solution for five minutes in triplicate. The passive conjugates, shown in Figure 4.40, produced a faint spot for the higher protein concentration, but was clearly inferior to Protein G-based conjugates discussed in Section 4.8.

As a result of the non-specific binding of the passive conjugates as well as the reliance on BSA as a component of the commercial dipstick assay blocking solution

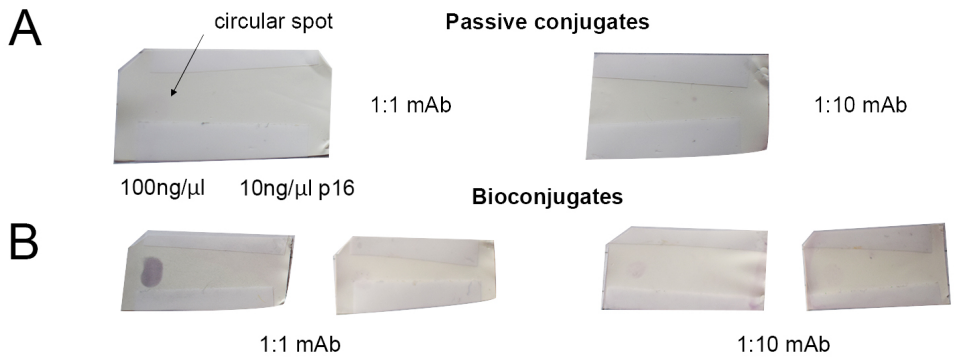


FIGURE 4.40: Dot blots carried out for A) passive conjugates at different antibody concentrations; B) bioconjugates (Protein G') with different antibody cconcentration.

for testing successful conjugation, it was decided to explore alternative methods for conjugation as seen in Section 4.8.

4.8 Bioconjugation of antibody

4.8.1 Bioconjugation scheme

As explained in Section 4.1, bioconjugation to AuNPs takes two major forms: native binding with Protein A or G and binding of biotinylated antibodies using avidin or streptavidin. The primary advantage of the Protein A/G method in particular is that by immobilising the antibody of interest onto a surface by its heavy chain within the Fc region, the functional region of the antibody is automatically oriented to offer optimal binding (Figure 4.4). Much like functionalised PEG it can also act as a blocking agent for the particle surface if titred properly. Given the issues discovered in Sections 4.6 and 4.7 with conjugate binding affinity in paper-based devices and the proven ability of Protein A/G mix to successfully bind the antibody of interest in a commercial dipstick assay, this approach provides an obvious alternative to the more common methods trialled in the remainder of the chapter.

4.8.2 Optimisation of Protein G' coverage

Protein G' (Proprietary Source, Recombinant, expressed in *E. coli*) was chosen as the bioconjugation moiety for the project owing to its ability to strongly bind murine immunoglobulins compared to Protein A. Recombinant Protein G' is a modified form of

Protein G in which the albumin and cell surface binding domains have been eliminated in order to reduce non-specific binding. The protein is supplied as a lyophilised compound and was reconstituted in phosphate buffer (pH7.6), and is nominally stable from pH2-10 with a pI of 4.5 and optimal binding pH of 5.0. In order to approximate the concentration order of magnitude required for a monolayer on the AuNP surface, the molecular footprint was calculated. Recombinant Protein G' is 21.6kDa but appears as 31-34kDa by Western Blot. Assuming a spherical molecule its diameter is then 4nm, giving a footprint of approximately $12nm^2$. The NP surface area is approximately $5050nm^2$. Thus, a monolayer requires 420 protein molecules per particle. Given an AuNP concentration of 16nM, the Protein G' solution should be 6.8µM. For a 21.6kDa protein this is equivalent to 3.67µg in 25µl or 0.15mg/ml.

BPE-AuNPs (0.75µM initial concentration) were produced as outlined previously and tested with resuspension in pH5.0 citrate buffer, which aggregated the particles completely. A 10mM solution of MES buffer was then produced and adjusted to pH4.8, in which the BPE-AuNPs remained stable without adverse effects on Raman signal (Section 4.4.4). Small volume tests (250µl) were produced according to Section 4.3.5 and a simple salt aggregation test was used to assess the coverage of Protein G on the surface. This was done by diluting the protein-coated particles with an equal volume of 1.71M NaCl solution and measuring subsequent aggregation optically - the absorbance characteristics before and after salt addition can be seen in Figure 4.41, with physical characteristics in Figure 4.42. These indicate that a minimum concentration between 0.038 and 0.075mg/ml is required to prevent aggregation in a strong ionic buffer, and that a concentration of 0.300mg/ml results in an excessive drop in absorbance. The sudden increase in zeta potential stability at 0.150mg/ml may indicate monolayer coverage, which corroborates the molecular footprint calculations seen above. Thus, the 0.075 and 0.150mg/ml sample protocols were carried forward for antibody titration.

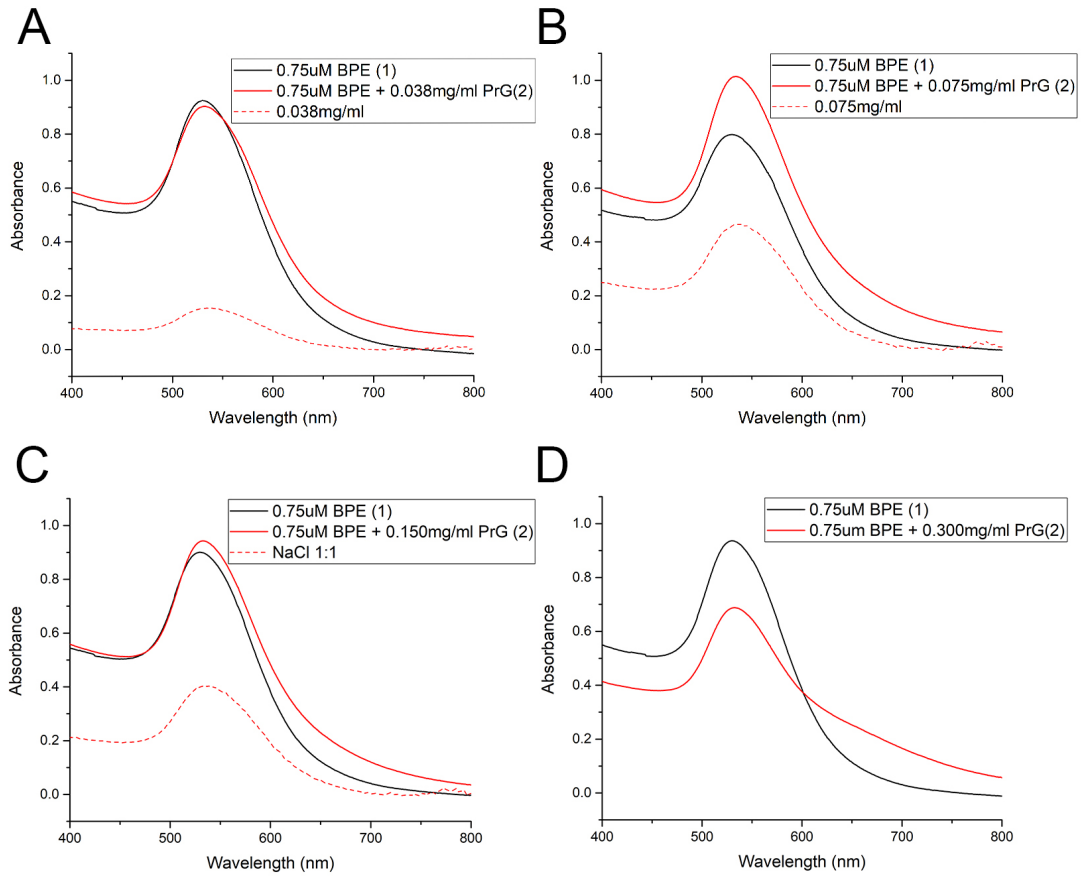


FIGURE 4.41: Extinction characteristics of BPE-NPs coated with Protein G' (solid red line) at initial concentrations of A) 0.038mg/ml; B) 0.075mg/ml; C) 0.150mg/ml; D) 0.300mg/ml compared to samples prior to conjugation (solid black line) and conjugated samples after salt aggregation (dotted red line).

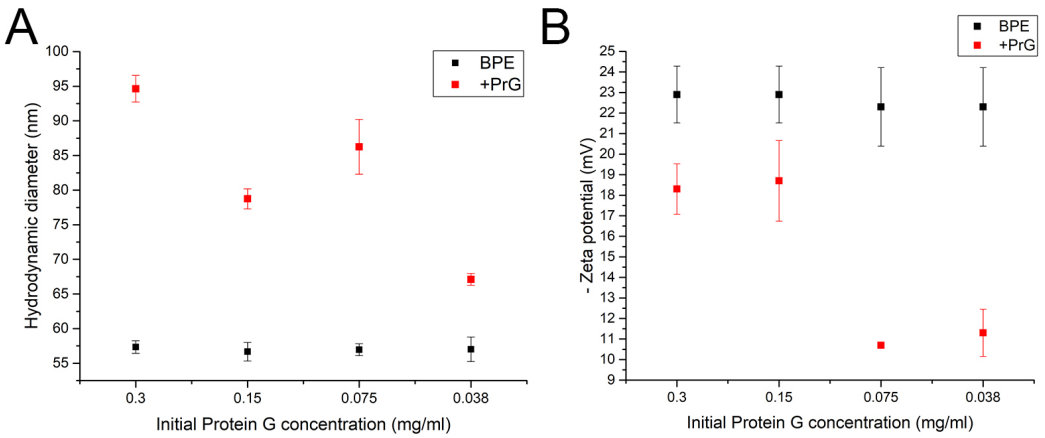


FIGURE 4.42: The A) hydrodynamic diameters and B) surface charge potentials of BPE-AuNPs before and after Protein G' adsorption at different concentrations. Protein concentrations 0.075mg/ml or lower result in a drop in zeta potential which was confirmed by salt aggregation.

4.8.3 Conjugation of antibody on Protein G′ coated particles

Undiluted antibody used with the initial protocol resulted in a positive dot blot signal on milk-blocked membrane (1% fatless milk powder in PBS for 30 minutes, incubation for two hours, washed for 5 minutes with milk solution in triplicate). Further samples produced without centrifugation and resuspension after addition of Protein G found a strongly reduced signal by dot blot, indicating that antibody was binding to excess Protein G in solution.

Further samples produced according to the specifications shown in Table 4.5 were also incubated on nitrocellulose membrane according to the same protocol. In each case 1µl spots of purified p16 protein (100ng/µl and 10ng/µl) were applied to a PBS-wetted nitrocellulose membrane and left to dry before blocking. The resulting colorimetric signals can be seen in Section B of Figure 4.40, which demonstrates clearly visible binding for all samples at different intensities. As expected, higher concentrations of antibody result in a greater colorimetric intensity, and so additional samples with 1:1 antibody dilution were produced and compared in Figure 4.43. Here the membrane was not wetted before use. The most successful combination of reagent concentrations (Sample 3.1, 0.150mg/ml Protein G′, 1:1 mAb) were reproduced in triplicate, tested with wetted membrane and found to be consistent. The conjugates were not found to bind non-specifically to high molecular weight proteins in a Western Blot format, as was the case with passive conjugates.

Previous experiments had shown that two hours was an optimal incubation time for conjugates on blocked membrane (when compared with 1, 3 and 4 hour incubations). The optimised sample was also applied to a commercial dipstick assay which had also been applied with 1µl spots of 100ng/µl and 1µg/µl and left to dry before conjugate wicking. This produced clear positive colorimetric signals after just ten minutes, indicating that the dipstick approach is preferable to surface contact only (also Figure 4.43). It was concluded that the two spots produce the same intensity of signal

Sample (500µl)	**50µl Protein G′ (**mg/ml**)**	**5µl mAb dilution**
1	0.075	1:1
2	0.075	1:10
3	0.150	1:1
4	0.150	1:10

TABLE 4.5: The concentrations of Protein G′ and antibody used for initial antibody conjugation experiments.

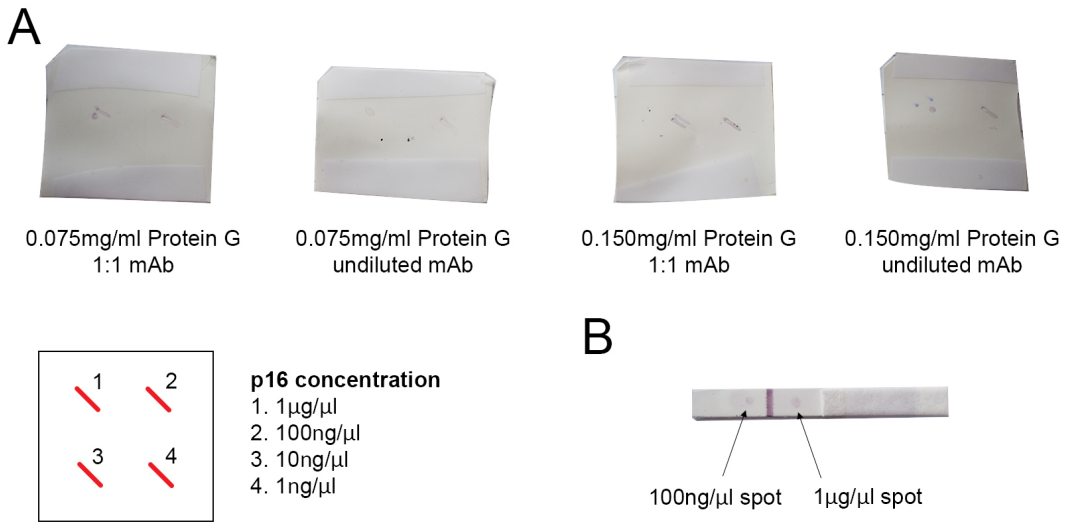


FIGURE 4.43: Colorimetric reponses of A) different bioconjugate conditions against purified p16 dot blots at different concentrations and B) the best canditate in dipstick assay format.

due to the expenditure of the majority of particles before the higher concentration spot was reached. It is unclear, however, why a positive test line is still seen at the Protein A/G line.

4.8.4 Storage and signal considerations

Having optimised a suitable protocol for conjugates which bind clearly and rapidly to purified p16 in a paper-based assay format, the bulk solution was interrogated with a Raman spectrometer (Section 4.3.2) and found to produce only small peaks, indicating that buffers used for the various conjugation stages and storage may need to be altered. As such, six samples were produced according to the specifications in Table 4.6.

Samples 2, 3 and 6 aggregated after a 48 hour period, indicating that HEPES is an inappropriate storage buffer for these conjugates. The remaining samples demonstrated, again, weak Raman signals (Figure 4.44). The stability of the conjugates overall indicated that phosphate buffer could be preferentially used as an antibody conjugation medium. As such, all subsequent conjugates were produced according to the following protocol:

Sample	mAb conjugation buffer	Storage buffer
1	MES (pH4.8)	Phosphate (pH7.6)
2	MES (pH4.8)	PBS
3	MES (pH4.8)	HEPES (pH8.0)
4	Phosphate (pH7.6)	Phosphate (pH7.6)
5	Phosphate (pH7.6)	PBS
6	Phosphate (pH7.6)	HEPES (pH8.0)

TABLE 4.6: The conjugation buffer conditions used with optimised concentrations of Protein G' and mAb.

1. Add 5µl desired reporter concentration to 495µl AuNPs and shake for 30 minutes at 500rpm.
2. Centrifuge at 2300g for 15 minutes to pellet. Remove 480µl supernatant and suspend in equal volume of 10mM pH4.8 MES buffer.
3. Add 50µl desired concentration of Protein G (mg/ml in pH7.6 phosphate buffer).
4. Centrifuge at 2300g for 15 minutes to pellet. Remove 480µl supernatant and suspend in equal volume of 10mM pH7.6 phosphate buffer.
5. Add 5µl desired antibody concentration to reporter AuNPs and shake for 30 minutes.
6. Centrifuge at 2300g for 15 minutes to pellet. Remove supernatant. Resuspend in either 480µl 10mM pH7.6 phosphate buffer for normal concentration or 80µl for a 5x optical density conjugate suitable for blotting.

It was found to be important that Protein G' in particular is added at the volume and concentration specified. A smaller volume of higher concentration providing the equivalent mass of protein (5µl 1.5mg/ml instead of 50µl 0.15mg/ml) was found to completely aggregate the nanoparticles.

A further set of conjugates investigating higher concentrations of reporter in conjunction with other reagents was produced as seen in Table 4.7. Their absorbance, Raman spectra and binding strength are shown in Figure 4.45. Samples 1, 5 and 8 produced the highest peaks, but Sample 8 also suffered from a high background. It is possible that this is due to partial aggregation caused by incomplete surface coverage.

A final set of BPE-AuNPs investigating higher concentrations of BPE suspended in MES buffer as though for bioconjugation are recorded in Figure 4.46. All were found

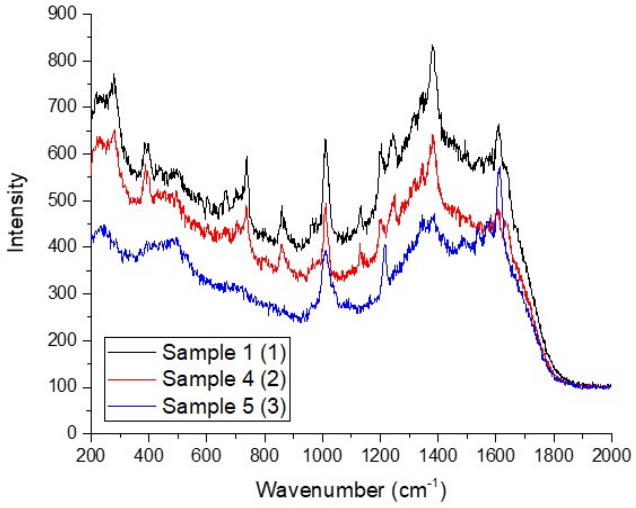


FIGURE 4.44: Raman spectral variation with conjugation and storage buffer of the three conditions in Table 4.6 which did not aggregate.

to be stable, demonstrating the importance of centrifugation and buffer conditions if compared to the stability of initial reporter experiments in Section 4.4.3.

A final set of optimised conjugates with increased BPE concentration were produced in triplicate with hydrodynamic diameters in Table 4.8, and absorbance and Raman spectra in Figure 4.47. These conjugates used initial concentrations of 5µM BPE, 0.15mg/ml PrG' and a p16 antibody dilution of 1:4, equivalent to 0.125mg/ml. The addition of protein layers reduces Raman intensity, although the signal produced for the particles in triplicate does vary for one sample. This conjugate method was used for all further experiments as seen in Chapter 5.

Sample	BPE (μM)	Protein G (mg/ml)	mAb dilution	Hydro. diameter (nm)
AuNP	-	-	-	48.0±0.3
1	1.25	0.150	1:1	124.8±4.6
2	0.75	0.150	1:1	123.2±1.8
3	1.25	0.075	1:1	204.3±48.3
4	0.75	0.075	1:1	133.4±4.7
5	1.25	0.150	1:4	122.1±7.7
6	0.75	0.150	1:4	287.2±48.1
7	1.25	0.075	1:4	250.4±37.0
8	0.75	0.075	1:4	133.0±3.4

TABLE 4.7: Further bioconjugate conditions investigating higher concentrations of BPE for improvement of Raman signal.

Sample	Hydrodynamic diameter (nm)
BPE Conjugate 1	130.4±4.1
BPE Conjugate 2	141.9±7.1
BPE Conjugate 3	146.4±6.5

TABLE 4.8: Hydrodynamic diameters of the final bioconjugate condition produced in triplicate.

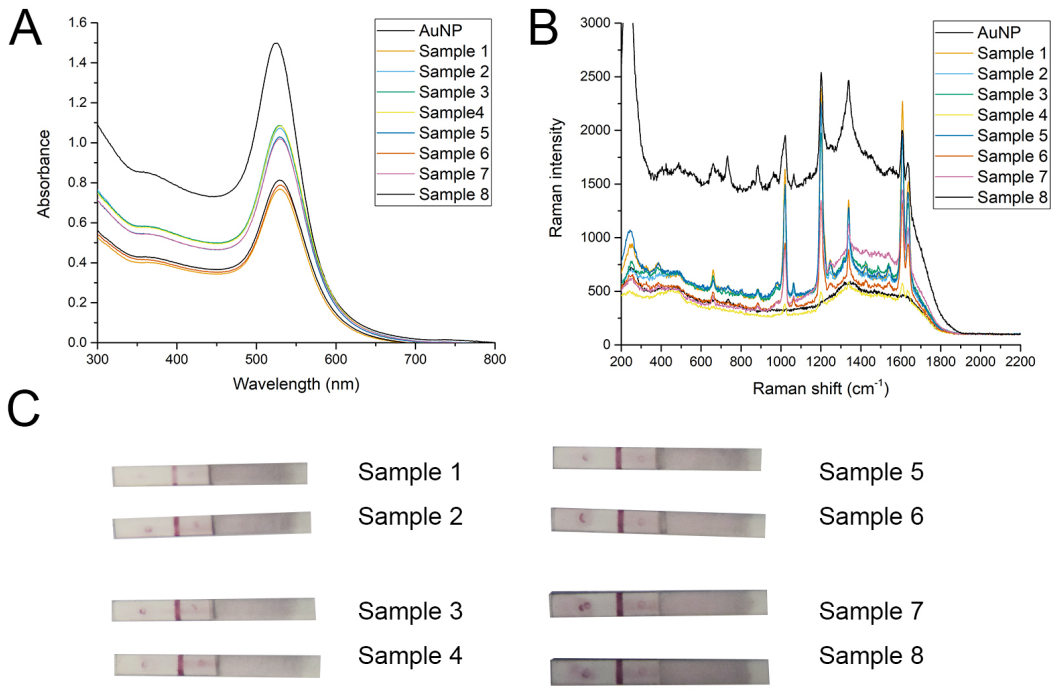


FIGURE 4.45: A) Extinction, B) Raman spectral and C) colorimetric binding characteristics of bioconjugate samples as detailed in Table 4.7.

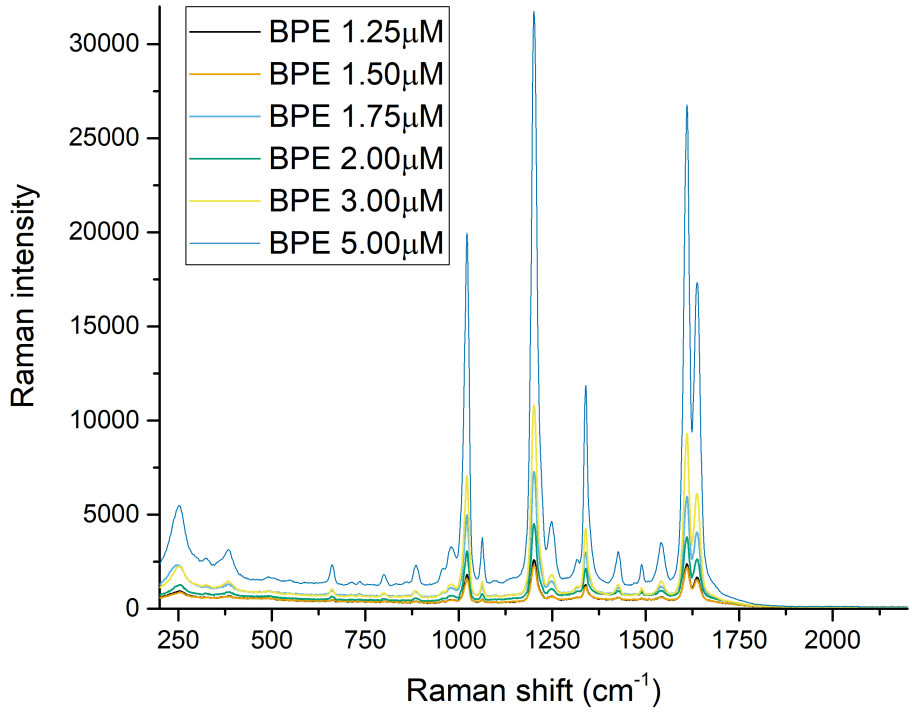


FIGURE 4.46: Raman spectra of higher concentration BPE-AuNPs without protein coverage (MES buffer).

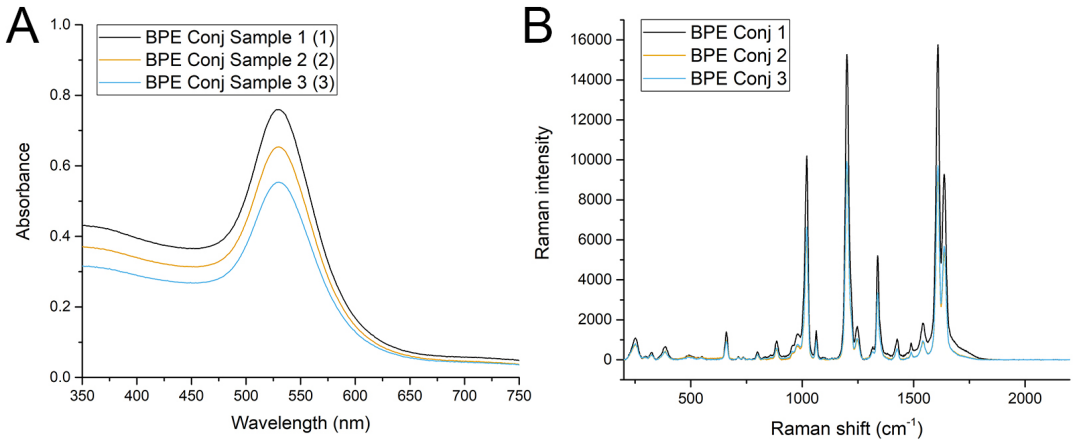


FIGURE 4.47: A) Extinction and B) Raman spectra of finalised bioconjugates produced in triplicate.

4.9 Conclusions

In this Chapter, a SERS-active antibody conjugate has been developed and optimised which binds to purified p16 protein. The individual components of an antibody-conjugated nanoparticle - metal substrate, Raman-active label, blocking moieties, binding ligand and conjugation method - have been characterised and optimised both independently and in conjunction with the other components. Three common conjugation methods were explored in order to produce the most effective conjugate for quantification of p16 levels in cervical cancer cells.

Covalent conjugation, although typically considered the 'gold standard' of conjugation approaches, demonstrated issues with non-reproducibility which have been documented previously and are most likely due to spontaneous cross-linking of the conjugation moieties and antibody. This leads to the demonstration of successful antibody conjugation onto the nanoparticle substrate but with disruption of antibody binding functionality. Passive conjugation provided useful information for the rapid optimisation of antibody concentration, but was found to produce weakly binding conjugates, most likely due to the random orientation of ligand which is intrinsic to the method. The Protein G bioconjugation method allowed not only for rapid optimisation but also for conjugates where bound antibodies are guaranteed to be oriented in a manner conducive to binding. The fact that the binding ligand acts as a protein block on the nanoparticle surface when titrated to monolayer coverage helped to reduce the number of conjugation steps as well as issues experienced with the use of BSA as blocking agent on passive conjugates. The final Protein G conjugate described can be produced in 2.5 hours at room temperature at a material cost of £2.32 per 1ml (unconcentrated), demonstrating baselined Raman intensity of 11530 $\pm$ 2650 counts without concentration and a clear colorimetric and Raman response to 50ng purified p16 in a paper-based assay.

An appreciation of the dynamics, competition and context of these critical components is clearly essential for production of an NP conjugate which is stable, reproducible and effective. Previously optimised covalent methods were ineffective for the antibody used, even when of the same isotype. Antibody sensitivity and the subsequent required concentration for labelling has a significant impact on substrate availability for Raman labels. Conjugation approaches may disrupt the biological functionality of the antibody being conjugated. Certain conjugation buffers have demonstrated capability to deteriorate SERS enhancement of Raman-active molecules. Spacers may improve binding site access and reduce non-specific interactions but also increase the size of the conjugate beyond usefulness. In order to capture the complexity of these

interactions for the production of an effective gold nanoparticle-based conjugate from scratch, a recommended general workflow has been produced and is shown in Figure 4.48.

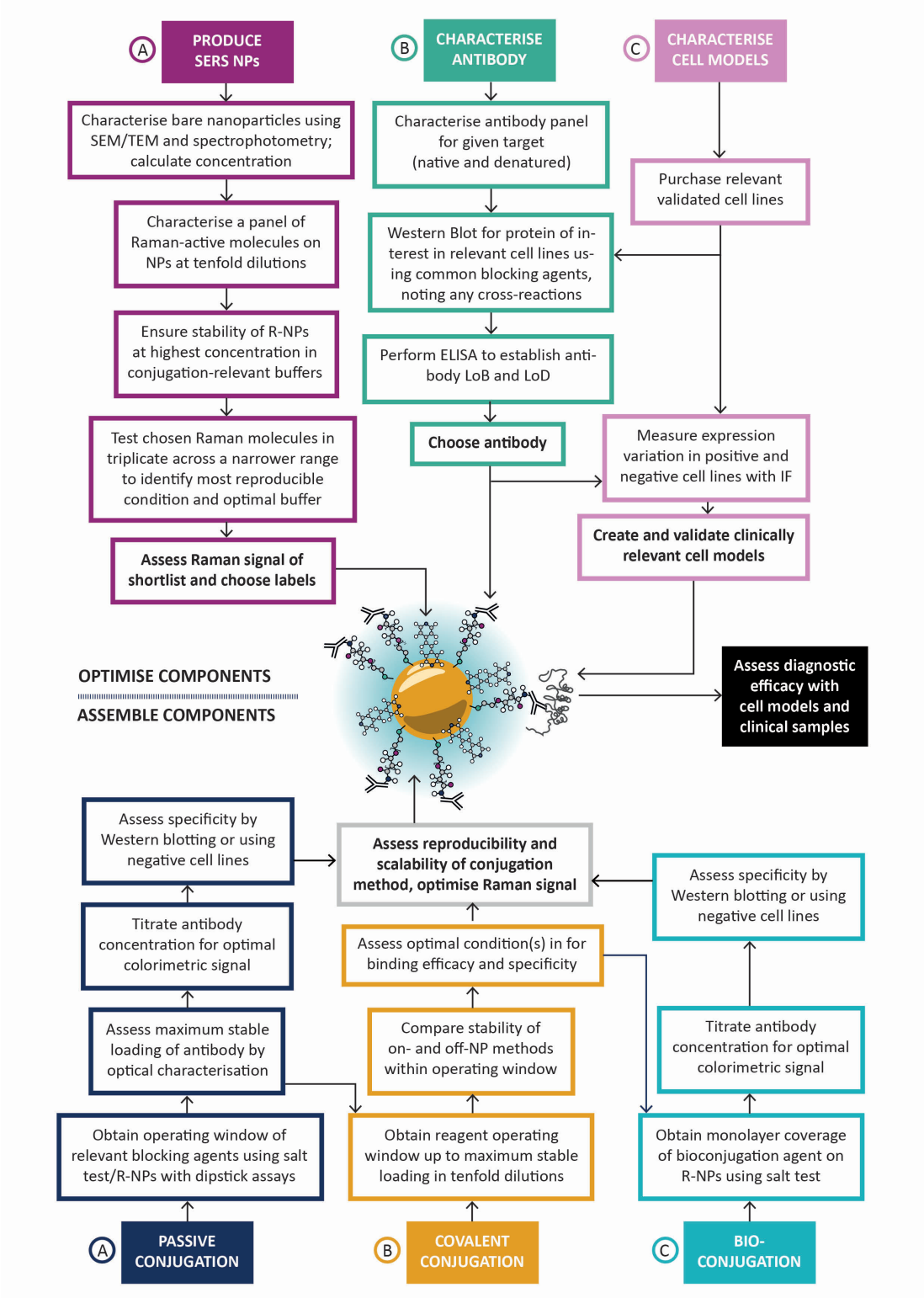


FIGURE 4.48: A recommended workflow for the development of SERS-active antibody-conjugated nanoparticles for use in diagnostic applications using three common methods of conjugation.

Chapter 5

Quantification of p16 in whole cells

5.1 Introduction

A targeted approach to molecular pathology allows for a diagnostically relevant marker to form the basis of a rapid, quantifiable test, rather than relying on unknown cellular changes as seen in Chapter 3. Using a targeted in-cell approach for cervical cytology would be useful as an aid for pathologists, as smear samples are phenotypically heterogeneous. Technically, a single highly abnormal cell in a sample of approximately 10,000 may be sufficient for a positive diagnosis, although in practice larger abnormal populations are required in order to be definitive.

A successful in-cell Raman diagnostic test has multiple requirements. The labelling moiety must of course be specific for its target, but should also be sufficiently stable for both consistent large-scale production and introduction into the cellular environment of choice without loss of structure or functionality. A protocol must be developed which minimises unwanted aggregation, residue or non-specific binding of the label. If multiplexing is used, the associated vibrational peaks must be suitably distinct for straightforward analysis, even in the presence of spectral background or particle aggregation. Finally, the label must be sufficiently well characterised to provide a confident and meaningful reference when quantifying target expression.

These requirements necessitate the development of a targeted imaging approach specific to each clinical application, including overall clinical goal, cell preparation and fixation state, expression location of the target marker and imaging setup. This is alongside the usual considerations of sample throughput and acquisition times.

This Chapter details the exploratory work performed in choosing a targeted imaging approach which is appropriate to the cervical cancer diagnostic context. Three approaches were investigated - conventional conjugate-based SERS imaging, SERS-SRS hybrid imaging and global-targeted hybrid imaging. The first two make use of the p16-specific nanoparticle conjugate produced in Chapter 4, whereas the third explores

the use of directly Raman-labelled antibodies with SRS rather than SERS enhancement mechanisms. Owing to limited data collection from the SERS-SRS method, this work is detailed in Appendix B (Chapter 9).

SERS imaging with nanoparticle conjugates

A vast number of biomedical studies have been undertaken with SERS-active nanoparticle conjugates, including live cell studies [197] not appropriate to this project - these are reviewed elsewhere [198, 199]. Those specifically in the cervical cancer context have investigated expression of human protein tyrosine kinase-7 (PTK-7) on HeLa cells with dual SERS-fluorescent nanorods [200]. PTK-7 has been demonstrated as a marker of cervical disease status and prognosis [201], but is not currently used in clinical settings.

Global-targeted hybrid imaging

Targeted stimulated Raman imaging is emerging as a potential competitor for fluorescence-based studies, owing to the small size of the labelling moiety (which is advantageous for live experiments) and the ability to simultaneously undertake label-free 'global' imaging of Raman-active groups in the sample which would provide complementary morphological and structural information [202, 168]. A number of publications have either developed multiplexed, highly Raman-active vibrational tags for SRS imaging [203] or demonstrated their use in drug discovery [204] and glucose uptake [205] amongst others.

5.2 Concept: ratiometric imaging approaches

While Raman-active small molecule antibody tags are conceptually similar to the use of radioactive or fluorescent labelling of antibodies for imaging purposes, the use of nanoparticle-based systems requires consideration of additional parameters. Nanoparticle conjugates are typically much larger (15-40$\times$) than alternative labelling methods, resulting in variations in uptake and retention of the probes dependent on the morphology of the sample rather than target marker expression. Studies by Wang et al. [206] indicated higher uptake of nanoparticle conjugates in normal *versus* xenograft tumour, whereas Kircher et al. [207] and others [208] have observed enhanced penetration and retention (EPR) in tumours, as one might expect.

Thus, a ratiometric system has been proposed previously in which a non-specific control or 'dummy' conjugate population is introduced alongside the target-specific

probe(s). This dummy conjugate has the same physical characteristics as the target conjugate but is biologically inactive and carries a Raman reporter with at least one distinguishing vibrational peak.

The level of target-specific binding can then be quantified once washing steps have been carried out to remove unbound probes. As the system provides pure reference spectra, direct classical least squares ('DCLS', also known as classical least squares, direct least squares or Beer's Law method) can be applied to analyse analyte concentration. Tested over large biological populations, this can eventually be linked to disease status in individual cells within a given smear sample.

DCLS assumes that at each wavenumber, the spectrum is the sum of a linear weighting of analyte reference spectra as shown in Equation 5.1. The algorithm iterates through spectral weightings until the sum of squared residual terms R is optimally reduced, where R is determined by the difference between the experimental and modelled value. The spectral weighting is proportional to the concentration of the analyte; owing to the multiple parameters which influence the proportionality constant, concentrations are compared ratiometrically between samples or determined using an internal standard [209]. As seen in Equation 5.1, accurate DCLS requires spectral preprocessing to remove cosmic rays and non-specific background contributions.

$$S = \Sigma_n w_n F_n + kB + \Sigma_m a_m P_m + R \tag{5.1}$$

where S is the raw spectral intensity, w_n and F_n are respectively the weighting and reference spectrum of nanoparticle type n, kB is the k-weighted average sample background spectrum B, a_m is the weighting of polynomial term P_m used for baseline correction and R is the residual term [210, 211].

5.3 Materials and methods

Sterilised phosphate-buffered saline (PBS) and ultrapure water ('dH_2O', MilliQ) was used for all experiments. All plate shaking steps were carried out at room temperature (RT) on a benchtop plate shaker (IKA MS 3) set to 500rpm. All optical components were supplied by Olympus unless otherwise stated.

5.3.1 Cell incubation with conjugates

Cells were seeded on 20mm diameter glass cover slips in six-well plates (Greiner Bio-One) and allowed to proliferate for 24-48hr at 37°C in 5% CO_2 until 30% confluent

for ease of individual cell imaging. For fixed cell incubation, growth media was aspirated, the wells washed three times with PBS and 500µl ice-cold fixing solution (either PreservCyt or 4% paraformaldehyde (PFA) in PBS) added for 30min. The fixative was aspirated, the wells washed again and cells blocked using 2% BSA in PBS for 1hr at RT. The blocking solution was replaced with 100µl 5× concentrated conjugates mixed with 400µl 0.1% BSA in PBS and the cells incubated for two hours at RT on a plate shaker. The conjugate solution was then aspirated and the cells washed twice with PBS and once with dH_2O, before being left to dry and sealed with clear varnish or mounted immediately using Fluorescence Mounting Medium (DAKO) onto glass slides depending on the experiment.

For live cell incubation, growth media was aspirated, the cells washed with PBS, 100µl 5× concentrated conjugates added to 400µl fresh media in each well and the cells incubrated at 37°C for two hours. The cells were then washed, fixed and mounted as described above.

All in-cell conjugate samples were imaged using an inVia Raman Microscope (Renishaw) with 785nm laser. Standard SERS mapping was used for several aspects of this thesis and so is described in Section 2.8, General Materials and Methods.

5.3.2 Fluorescent antibody labelling

Unlabelled whole goat anti-mouse antibody (CAT A28174), 1mg/ml in PBS + 0.09% sodium azide, was conjugated to Alexa Fluor 488 NHS ester (CAT A20000), succinimidyl ester, 1mg, molecular weight 643.4g/mol. This was compared to commercial Alexa Fluor 488 goat anti-mouse (CAT A11001), 2mg/ml in PBS (pH7.5) + 5mM sodium azide. For immunofluorescence, both secondary antibodies were bound to p16 INK4A monoclonal mouse anti-human antibody (BD Pharmingen), 0.5mg/ml in PBS.

The following protocol represents a slight variation of that described in 'Amine Reactive Probes' (Molecular Probes, Life Technologies). An Antibody Concentration Kit (ab102778, Abcam) was used to increase the concentration of the unlabelled antibody from 1 to 5mg/ml and simultaneously buffer-exchange the protein into 0.1M sodium bicarbonate buffer (pH8.3) for conjugation. Between 1.50 and 3.00µl amine-reactive fluorophore (5 or 10mg/ml) was slowly added to 100µl concentrated unlabelled antibody whilst vortexing. The tube was wrapped in foil and incubated for 1hr at room temperature with continuous stirring on a rotary incubator. The reaction was neutralised with 10µl freshly prepared hydroxylamine (pH8.5, 1.5M and incubated for 1hr as before.

To separate the conjugate from unreacted fluorophore, 500µl Sephadex-50 gel per sample was prepared in PBS and left to hydrate overnight. The gel slurry was added to a prepared filtration column and excess buffer left to run before the reaction volume was added. The first fluorescent band to elute is the successfully conjugated antibody. The absorbance spectrum of a 1µl volume of the sample was assessed, with values at 280nm and the absorbance maximum of the dye (495nm for Alexa Fluor 488) recorded for degree of labelling (DOL) calculations, the equation for which is given below:

$$DOL = \frac{A_{max} \times MW}{[protein] \times \varepsilon_{dye}} \tag{5.2}$$

where A_{max} is the absorbance maximum of the conjugated fluorophore, MW is the molecular weight of the protein, ε_{dye} is the extinction coefficient of the dye at its absorbance maximum, and [protein] is the protein concentration in mg/ml. For IgG antibodies, 1.4 $A_{protein}$ units is equivalent to 1mg/ml protein concentration. The absorbance $A_{protein}$ of the antibody is calculated from

$$A_{protein} = A_{280nm} - (A_{max} \times CF) \tag{5.3}$$

where CF is the correction factor of the dye contribution at 280nm, given by $\frac{A_{280freedye}}{A_{maxfreedye}}$ and quoted as 0.11 for Alexa Fluor 488.

The labelled antibody was stored in the dark at 4°C and -20°C for the short (less than three weeks) and long term respectively.

5.4 Production of a physicochemical control particle

An isotype control or 'dummy' conjugate, which complements the finalised bioconjugate discussed in Chapter 4, was developed in order to assess the non-specific behaviour of nanoparticles in fixed cells. Initially, PYOT particles were produced due to the similarities in hydrodynamic size and surface charge found in comparison with BPE-coated AuNPs. However, the effect of such characteristics on a protein-coated particle are negligible in comparison to the need for clearly discernible Raman peaks. Two of the primary peaks for BPE and PYOT occur in the same region (1603-1606cm^{-1}) and so PPY was explored as an alternative.

5.4.1 Binding validation

Prior to use in cells, the specificity and binding affinity of target and dummy conjugates was tested using lateral flow assays. Paired conjugates were initially produced

with parameters outlined in Table 5.1 and with the Protein G' conjugation method described in Section 4.8.4. Following the final centrifugation, 15µl of each conjugate was used diluted 1:1 with conjugate buffer (with BSA), equivalent to a final concentration of 25×. The low volume of conjugate required an additional 'flush through' with 15µl conjugate buffer to prevent nanoparticles from becoming embedded in the membrane prematurely. The experiment was repeated with 5x concentrated sample which provided a clear colorimetric response for target particles with 25ng purified p16 protein. The experiment was then repeated with optimised concentrations of BPE and PPY as discussed in Section 5.4.2, again demonstrating, as expected, a clear colorimetric response from the target particles and blank response from the dummy particles. Both tests can be seen in Figure 5.1.

Sample	Reporter	Antibody
Target Conjugate 1	-	1% 1:1 mAb
Dummy Conjugate 1	-	1% 1:2 mAb
Target Conjugate 1	1% 5µM	1% 1:4 mAb
Dummy Conjugate 2	1% 20µM	1% 1:8 mAb

TABLE 5.1: The reagent parameters used for preliminary and final ratiometric conjugates. Volumes are given as a percentage of the initial sample volume, e.g. 10% refers to 50µl reagent added to a 500µl sample. All samples used 10% 0.15mg/ml Protein G' solution for antibody conjugation. Binding validation is shown in Figure 5.1.

In both cases the isotype control 'dummy' particles did not produce a colorimetric response other than a 'coffee ring' effect for the high concentration sample with high protein content. This is posited as being due to impeded flow through the membrane - the effect is seen for both types of particles and indicates that the non-specific responses of the conjugates are similar in this format.

5.4.2 Physical characteristics

As with BPE titration, a range of PPY inital concentrations from 10^{-3} to 10^{-6}M were tested through 1:100 dilution with AuNPs. Concentrations 10^{-3} to 10^{-4} aggregated immediately. Further titrations of PPY at 10, 15, 20, 30, 40 and 50µM were characterised by extinction and Raman activity as shown in Figure 5.2: samples degraded in stability from 30µM onwards, producing large Raman peaks at 952cm^{-1} for the 40µM sample as a result of aggregate hotspot formation.

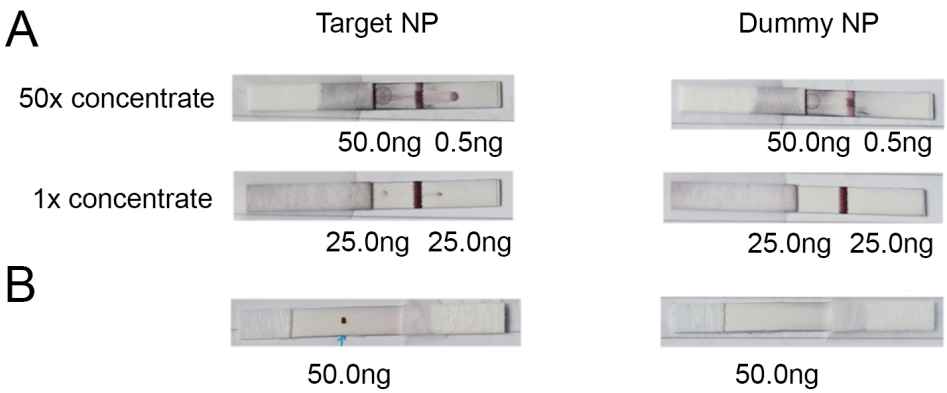


FIGURE 5.1: Binding validation of target and dummy conjugates by dipstick assay with purified p16 protein. A) 50× and 1× concentrated target and dummy conjugates without Raman reporters. B) The final conjugate designs tested at 5× production concentration. Both were tested against 0.5µl spotted volumes of the protein content shown. As expected, the dummy conjugate with isotype antibody control does not bind to the target.

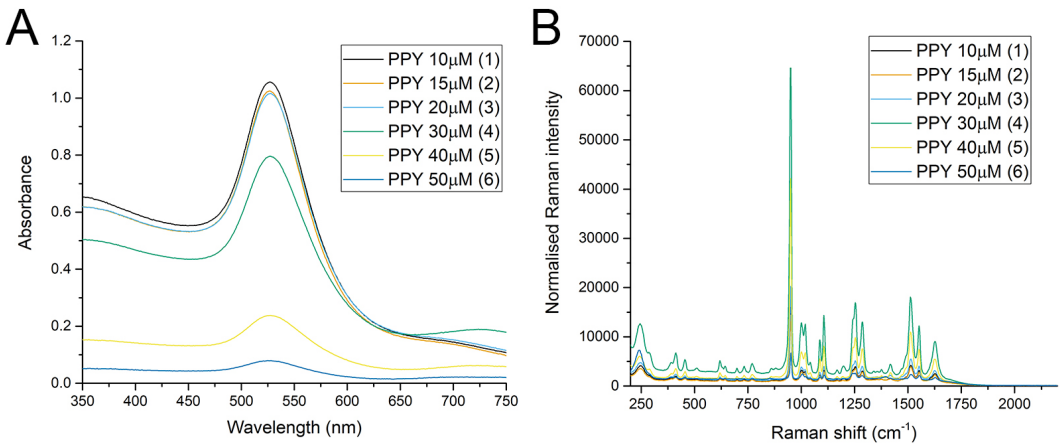


FIGURE 5.2: The absorbance and Raman spectra of bioconjugates produced with varying concentrations of PPY stock, used as a 1% volume of the sample. A) The colloidal solution is stable to 20µM, beginning to aggregate at 30µM and beyond. B) The conjugates show high SERS activity at unstable concentrations due to aggregation-induced hotspot junctions, but the stable concentrations still produce a high Raman signal at over 10000 counts for the primary peak.

PPY conjugates were then produced in triplicate using identical parameters to the BPE target particles, but using an initial concentration of 20µM. The extinction spectra can be seen in Figure 5.3, with physical characteristics summarised in Table 5.2.

The variation in ratio of extinction peak to 650nm, a concentration-independent indicator of stability, indicates variation in the PPY triplicate samples which may suggest that 20µM PPY may not be fully stable. Smaller ratios correspond to less pronounced absorbance peaks and therefore indicate aggregation.

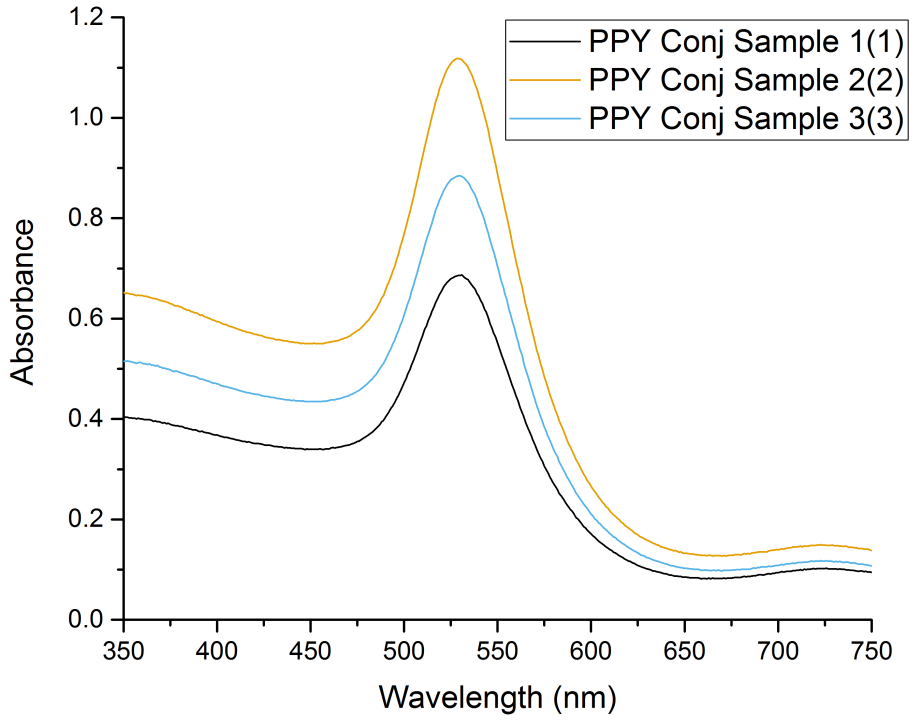


FIGURE 5.3: Absorbance characteristics of PPY samples in triplicate. The variation is partly caused by concentration variation but also due to reduced stability compared to the BPE target conjugates (see Table 5.2).

Target and dummy conjugates were mixed in bulk at ratios from 0:5 to 5:0 in 20% [v/v] increments and the combined Raman signals interrogated using 1.0s acquisitions with a benchtop Raman spectrometer (Snowy Instruments) with 100mW 785nm laser. The use of a 638nm laser line was found to significantly reduce Raman intensity. The results can be seen in Figure 5.4.

Figure 5.4A shows a clear but non-linear increase of both of the unique principal peaks for BPE (1201cm^{-1}) and PPY (952cm^{-1}) with relative volume. Figure 5.4B, however, shows that analysis can not simply be made on the basis of peak height. Fractional variations in concentration can produce a significant effect on the proximity of

Sample	A_{Peak}	A_{650}	Ratio
BPE 1	0.7596	0.0721	10.55
BPE 2	0.6538	0.0611	10.70
BPE 3	0.5533	0.0557	9.93
PPY 1	0.6863	0.0849	5.17
PPY 2	1.1180	0.1327	8.43
PPY 3	0.8846	0.1024	8.64

TABLE 5.2: Absorbance characteristics of target and dummy conjugates. The ratio of absorbance at the peak versus 650nm gives an indication of the aggregation status of the sample, with higher numbers indicating higher stability.

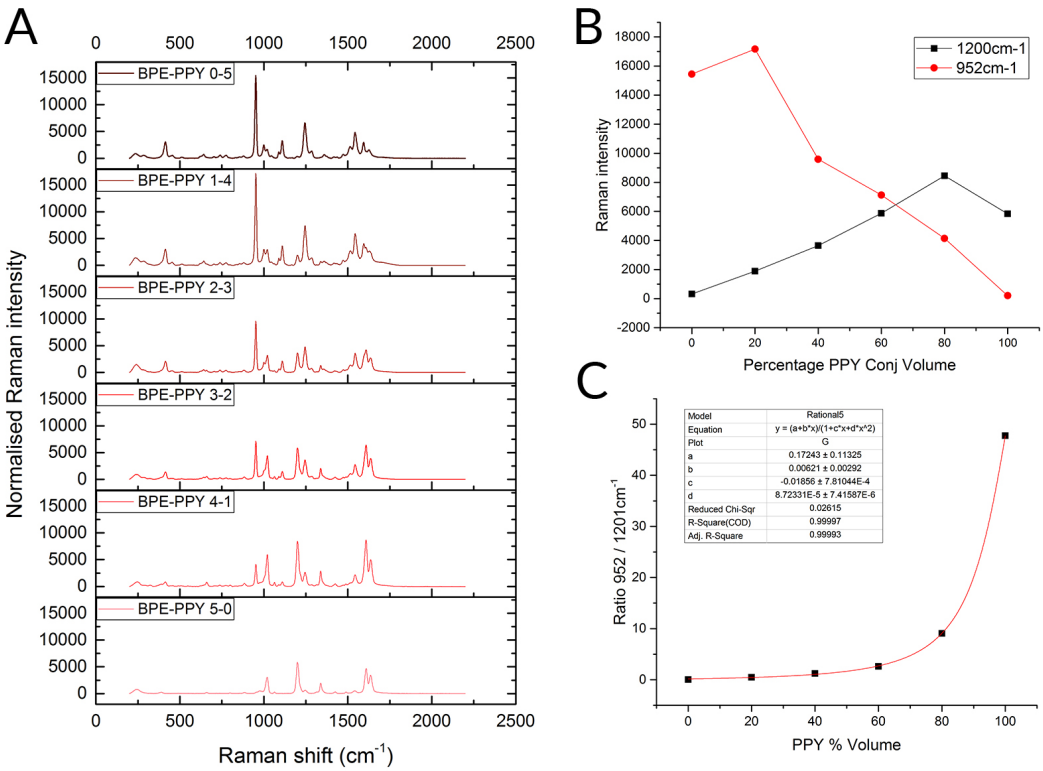


FIGURE 5.4: Reference ladder of bulk ratiometric samples, mixed at 20% intervals. A) The normalised Raman spectra of each sample. B) Intensity of principal peaks from each sample showing some concentration-dependent error. C) A ratiometric reference curve of the data, independent of concentration and showing a clear reciprocal trend.

particles within the acquisition volume and therefore the SERS signal. However, by taking a ratio of the unique peaks for the two molecules and plotting against relative

conjugate volume within the sample, a relationship can be found and potentially used to predict ratios of target and dummy conjugates within a sampling volume (Figure 5.4C). A 'Rational5' Nonlinear Curve Fit was used in Origin, to reflect the ratiometric (rational fraction) nature of the data. The potential instability of the 20µM PPY conjugate, combined with the steep slope of the ratiometric curve, suggest that a lower concentration of the dummy reporter may be beneficial to reduce potential errors in imaging data interpolation.

5.5 Ratiometric conjugates in whole cells

Having produced a complementary dummy particle for the established conjugate to p16 protein, the ratiometric system was explored with SERS mapping in fixed cells to test its ability to distinguish between expression levels in cell models and quantify the relative concentrations of nanoparticles present.

5.5.1 Preliminary optimisation of conjugate incubation

Initially cell samples were prepared according to the protocol described in Section 5.3.1 and left to dry before inverting. Target (BPE 5µM) and dummy (PPY 20µM) conjugate incubations were carried out with fixed HaCaT and SiHa cells, where p16 is respectively undetectable and overexpressed. Cells were fixed in either PFA (cross-linking) or PreservCyt (dehydrating) as a comparator of the standard approach against a clinically relevant model.

SERS mapping was carried out as described in Section 2.8 and mapping analysis conducted with component-based DCLS using bulk reference spectra. Representative imaging maps and associated spectra are shown in Figure 5.5. Owing to the use of drying for cell preparation, issues were encountered with salt crystal formation that made cell membranes difficult to distinguish, particularly for PFA-fixed cells. Table 5.3 shows the location of targeted conjugate hotspots, with signals designated as True Positive or False Negative for inside or outside SiHa respectively. This labelling convention was inverted for HaCaT cells. The SiHa and HaCaT cells incubated with dummy conjugates displayed hotspots inside and outside cells.

Overall the results indicate that the use of PFA as a fixative results in lower specificity of measurement, as all of the HaCaT cells prepared in this way show false positive results. This is posited as being due to the cross-linking effects of PFA, which are more likely to 'trap' conjugates non-specifically than cell dehydration. It is worth noting here that for two of the HaCaT PreservCyt samples, a small Raman signal was

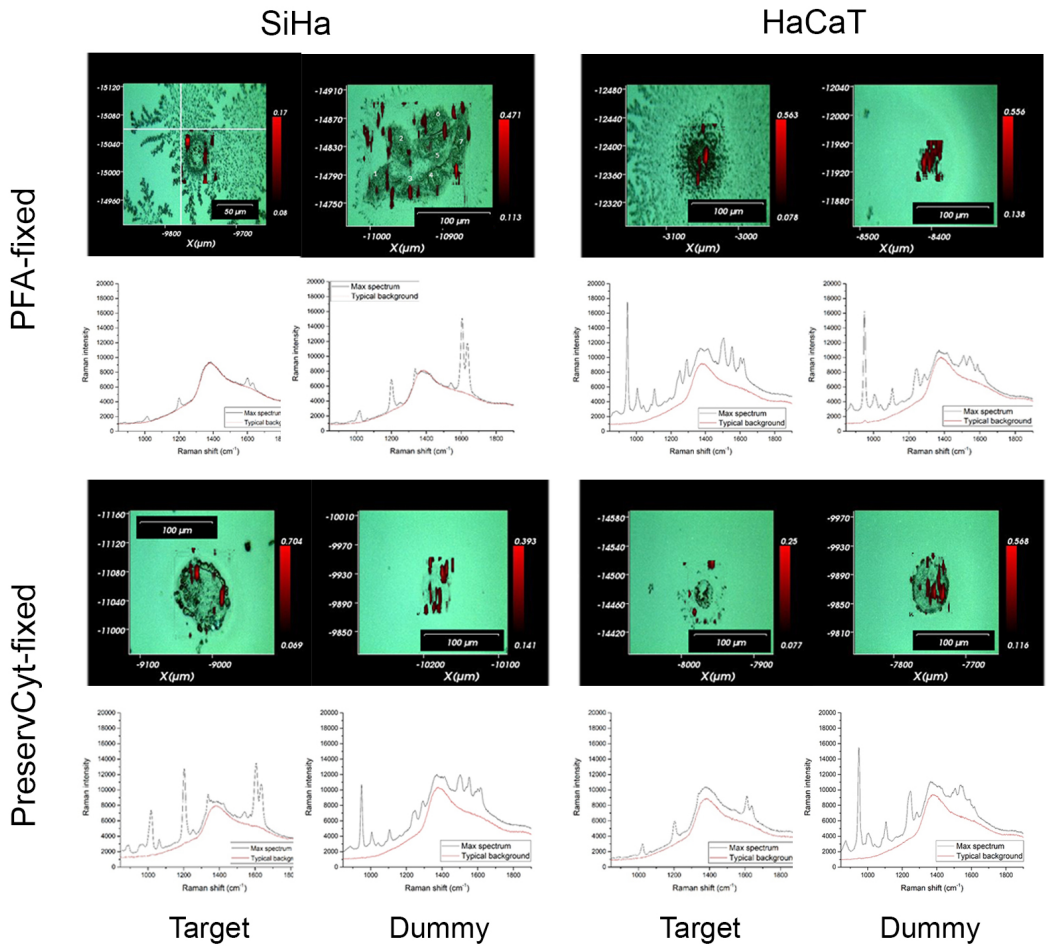


FIGURE 5.5: Examples of SERS maps and corresponding hotspot spectra for PFA- and PreservCyt-fixed wildtype SiHa and HaCaT cells, generated using Wire 4.2.

present barely visible above the fluorescent background. Use of this technique with clinical samples would require careful thresholding to determine the level of SNR required to constitute a signal. Finally, the use of morphologically distinct cell models, though appropriate in terms of p16 expression, do not give a comparable indication of non-specific accumulation and general kinetics of conjugates inside fixed cells.

Conjugate incubation with fixed cells in suspension was also tested. For this, 100µl aliquots of fixed HaCaT (p16 not detectable) and SiHa (p16 overexpressed) cells were washed with PBS as described in Section 3.3.1 to remove PreservCyt residue, before being blocked with 2% BSA in PBS for 1hr at room temperature (RT). Blocking solution was removed by pelleting the cells at 6800g for 5min and removing the supernatant.

Sample	Fixative	True	False	Sample	Fixative	True	False
SiHa 1	PFA	1	-	HaCaT 1	PFA	-	3
SiHa 2	PFA	3	-	HaCaT 2	PFA	-	4
SiHa 3	PFA	6	-	HaCaT 3	PFA	-	4
SiHa 4	PFA	-	1	HaCaT 4	PFA	-	8
SiHa 1	PreservCyt	-	2	HaCaT 1	PreservCyt	-	2
SiHa 2	PreservCyt	2	-	HaCaT 2	PreservCyt	0	-
SiHa 3	PreservCyt	4	-	HaCaT 3	PreservCyt	1	-
SiHa 4	PreservCyt	6	-	HaCaT 4	PreservCyt	2	-

TABLE 5.3: Labelling accuracy of target conjugates inside SiHa and HaCaT cells fixed with either PFA or PreservCyt. 'True' and 'false' designations correspond to True Positive and False Negative for SiHa (p16 expressed), and True Negative and False Positive for HaCaT (p16 not expressed).

Cell samples were incubated with 100µl 50× concentrated target or dummy conjugates for 1.5hr on a plate shaker. The cells were pelleted again and the supernatant, containing free conjugates, removed. The cells were suspended in 40µl PBS and divided between two cavity slides for imaging under the same conditions as described previously. This resulted in poor signal and issues with cell drift under laser interrogation due to lack of adherence.

On the basis of the experiments described above, several changes were made to the method. First, the mounting medium used for immunofluorescent slide preparation was tested and found to be vibrationally silent. Thus, slides were then prepared according to the immunofluorescence protocol described in Section 2.4 but using conjugate incubation instead of primary and secondary antibodies. Second, the dummy conjugate reporter concentration was reduced from 20µl to 10µl to reduce likelihood of aggregation. Finally, comparisons were made between SiHa w/t and transfected with p16 in order to study SERS signal without complications from differing cell morphologies.

5.5.2 Development of conjugate incubation and ratiometric method

A new set of conjugates was produced and incubated with SiHa cells in keeping with the changes described above. SiHa cells were transfected with his-tagged p16 as described in Section 2.2 for the highly expressing model. Initially, the conjugates were tested separately so that minimal signal processing was required. DCLS was used with the bulk reference spectra seen in Figure 5.6 to produce the false-colour hotspot maps seen in Figure 5.7 - the corresponding raw spectra are in Figure 5.8. The maps

themselves do not provide any visual indication of p16 over-expression. However, the spectral data (and corresponding look-up values for the map) provide a clearer picture. As hoped, the dummy conjugate signals vary slightly between samples but not significantly between w/t and +p16 SiHa. The peak heights of the target conjugates show approximately a 2 to 5-fold increase in signal for the +p16 compared to w/t SiHa.

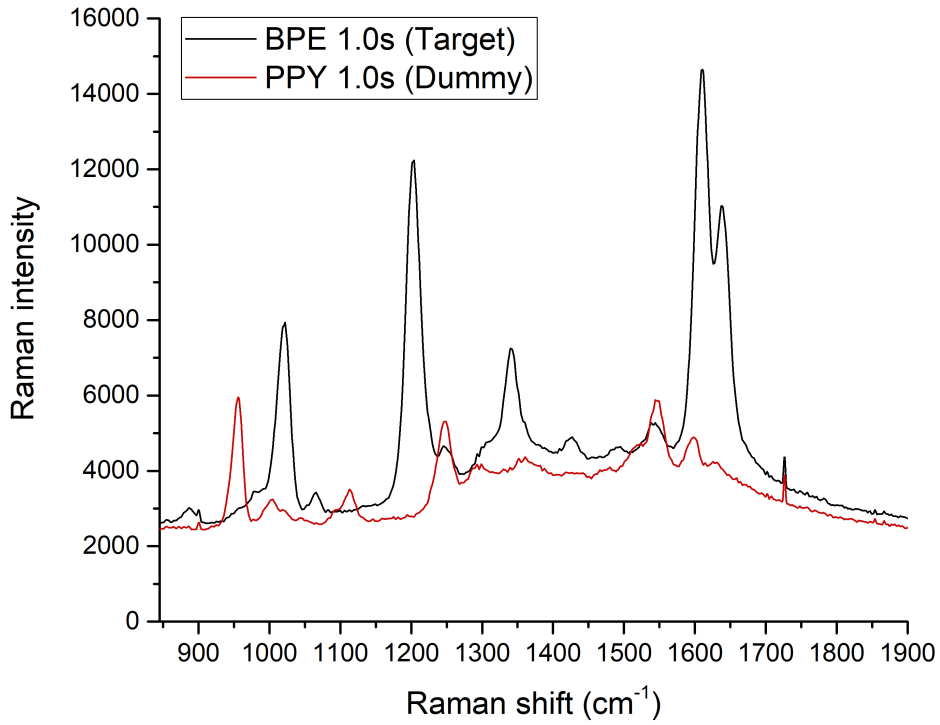


FIGURE 5.6: Bulk reference spectra for target and dummy conjugates, used in Section 5.5.2. Spectra were acquired on a CBEx 785 hand-held Raman identification system.

For ratiometric mapping, a second set of conjugates were produced with the same concentrations of reporter and incubated as a 1:1 volume mix with w/t and +p16 transfected SiHa, fixed with PreservCyt before incubation. Individual incubations were also produced. The cell maps can be seen in Figure 5.9, with spectra in Figure 5.10. The elongated nature of the hotspots is an artefact of the use of a line profile for the laser without a pinhole; this setup was found to significantly improve detectable signal. The white arrows indicate the pixels with maximum spectral intensity - for SiHa w/t, these coincide, but for SiHa +p16 they are distinct. Most importantly, the target

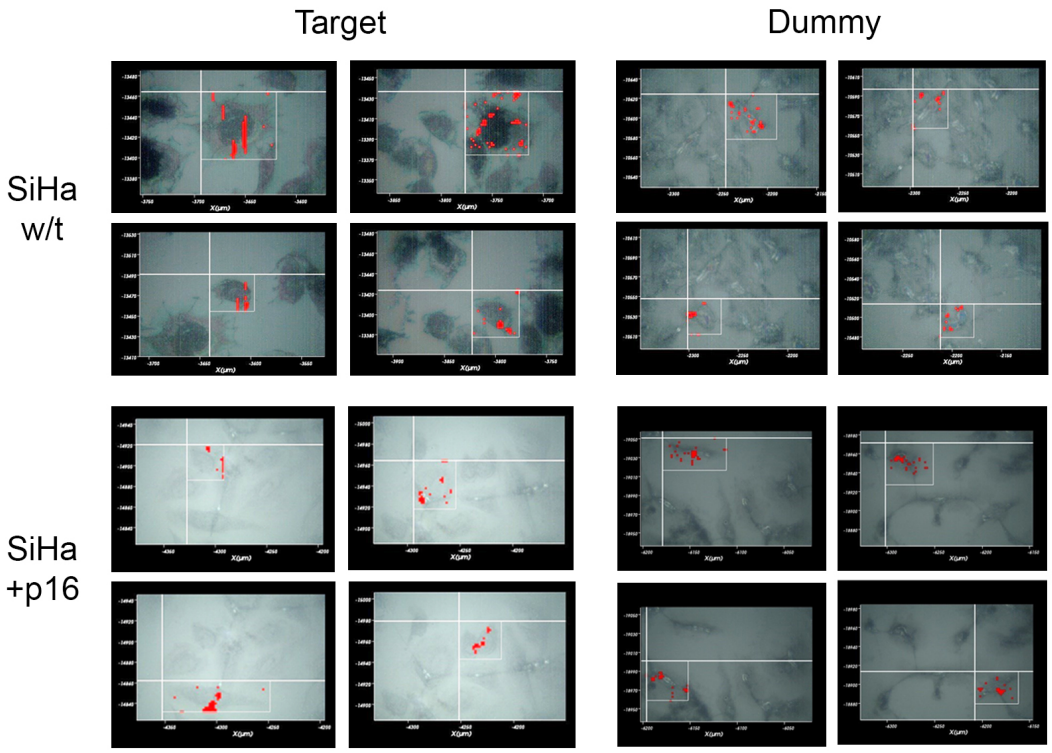


FIGURE 5.7: Examples of SERS maps and corresponding hotspot spectra for w/t and transfected PreservCyt-fixed SiHa, which do not give any visual indication of a link between target signal and p16 expression.

conjugate intensity varies in keeping with the expected expression of p16 in the cells, demonstrating a 2.2-fold average increase in 1201cm^{-1} BPE signal for the transfected model, whereas the dummy signal decreased by 1.2-fold.

For analysis of conjugate ratios, a reference ladder of 5µl mixtures from 100% to 0% target in 10% increments was spotted down onto a glass microscope slide and left to dry before imaging at 20× magnification. The two pure references were used to create the DCLS component maps seen in Figure 5.9, but it was noted that the dummy conjugate areas marked as high intensity by the DCLS analysis were not all genuine hotspots when compared manually. This was partly due to the presence of the fluorescent background which could not be removed through automatic baselining or substraction of a reference spectrum without altering central peak heights. The primary issue was that the principal PPY peak had deteriorated in the dried reference ladder. As seen in Figure 5.11, the peak is convoluted with small peaks in the BPE spectrum. Given the highlighted importance of signal quantification for progress to

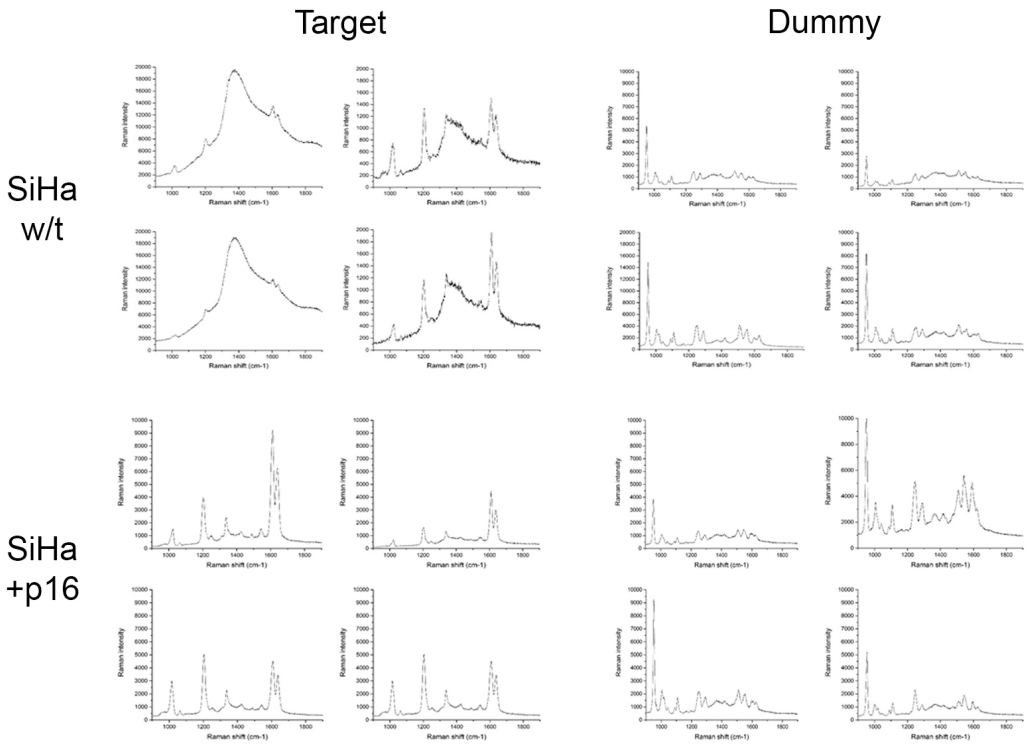


FIGURE 5.8: Raw hotspot spectral data for w/t and transfected PreservCyt-fixed SiHa, corresponding to the equivalent maps in Figure 5.7.

be made in SERS mapping and the successful demonstration of conjugates presenting in the expected levels in fixed cervical cancer cells, it was decided that signal processing of the spectral data should be prioritised.

5.5.3 Establishing a signal processing technique

Having established that a SERS signal could be detected from conjugates inside cells, efforts were focused on SERS signal processing and quantification. Dummy samples produced at 20µM PPY initial concentration were revisited to account for the signal crosstalk experienced with dried reference samples at 10µM PPY. A second ratiometric reference ladder produced and characterised in bulk and via mapping. The bulk ratiometric ladder consisted of target and dummy conjugates mixed in 25% increments to give 100:0, 75:25, 50:50, 25:75 and 0:100 ratios. Once bulk spectra had been acquired, 5µl aliquots of each sample were transferred to a glass slide to dry and then interrogated via Raman mapping. Bulk spectra were acquired over a single 1.00s acquisition

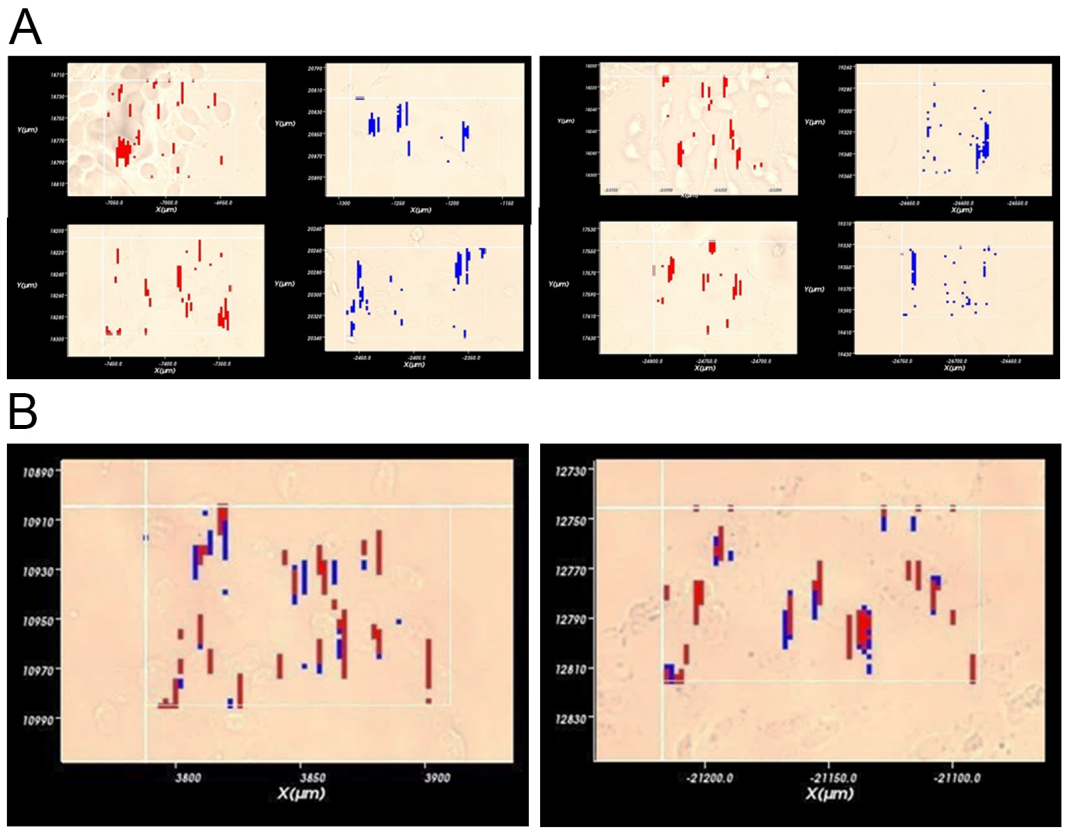


FIGURE 5.9: SERS hotspot maps of w/t (left) and p16-transfected (right) PreservCyt-fixed SiHa cells incubated with target (red) and dummy (blue) conjugates. A) Separate incubations, which, as before, do not give any visual indication of p16 expression status. B) Ratiometric incubation.

and mapped spectra acquired with HRStreamline on Wire 4.2 using a 2.00s acquisition and no pinhole. The normalised bulk spectra and associated peak height-based ratiometric reference curves can be seen in Figure 5.12. For the mapped data, 9-point spectral smoothing (third-order polynomial) was carried out and an automated baseline subtracted. In initial tests, each spectrum was truncated to the region of interest ((850-1250cm^{-1})) to remove the fluorescent background, but this process was found to crash the analysis software during mapping.

The processed spectral maps were then analysed ratiometrically using an in-house MATLAB script, modified from code written by Catherine Lynch at the University of Edinburgh and Lauren Jamieson at the University of Strathclyde. The commercially available software, Wire 4.2, was not capable of either direct ratiometric mapping or exporting matrices of DCLS weightings for further analysis.

The mapping script allows for the ratiometric value of two peaks within the same spectrum to be displayed as a false-colour pixel, allowing ratiometric maps to be generated. The resulting matrix of ratios corresponds directly to the percentage volume data shown for the conjugates in Figure 5.12C and D.

Figure 5.13 shows maps of the dried reference spots and a visual representation of the signal processing work flow. As the references progress through to 100% target, the number of red pixels (corresponding to high percentages of target) increases, whereas blue (100% dummy) and black (approximately 1:1 mix) pixel numbers are reduced. The granular nature of the reference data is due to the resolution of the technique, which can pick up locally dried areas of target or dummy conjugates not seen with bulk imaging. This actually gives a better indication of what spectral processing might look like for incubated cell samples. Calibration using the reference curve would provide a more linear response for the false-colour map.

The preliminary data shown stands as proof of principle that the ratiometric approach could be used incorporated into established software platforms in an automated fashion. This would give more meaningful data than conventional SERS maps, and could be applied to a fixed cell context.

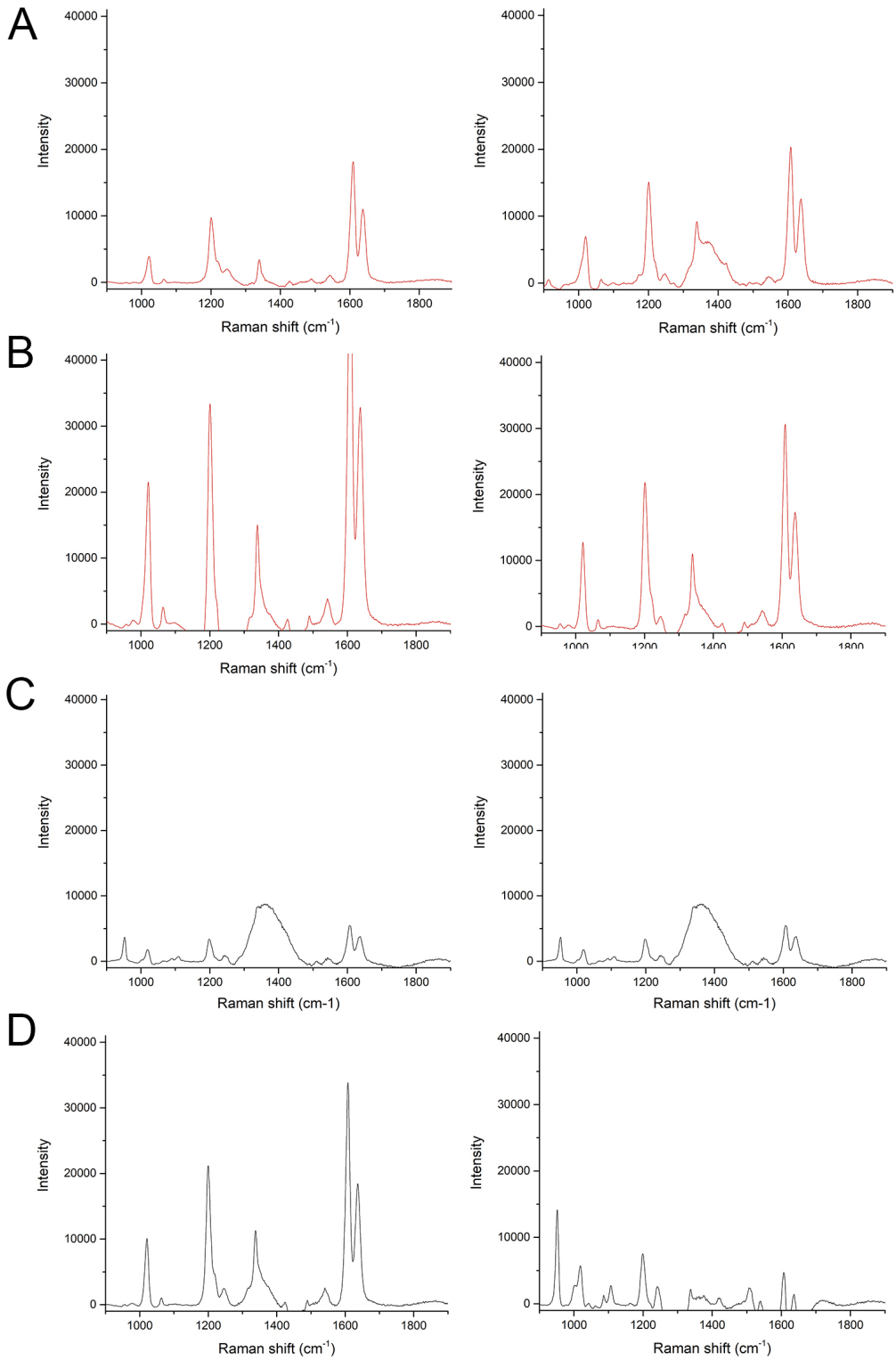


FIGURE 5.10: The baselined spectra from separate incubation SERS maps in Figure 5.9A. A) SiHa w/t target hotspot spectra. B) SiHa +p16 target hotspot spectra. C) SiHa w/t dummy hotspot spectra. D) SiHa +p16 dummy hotspot spectra.

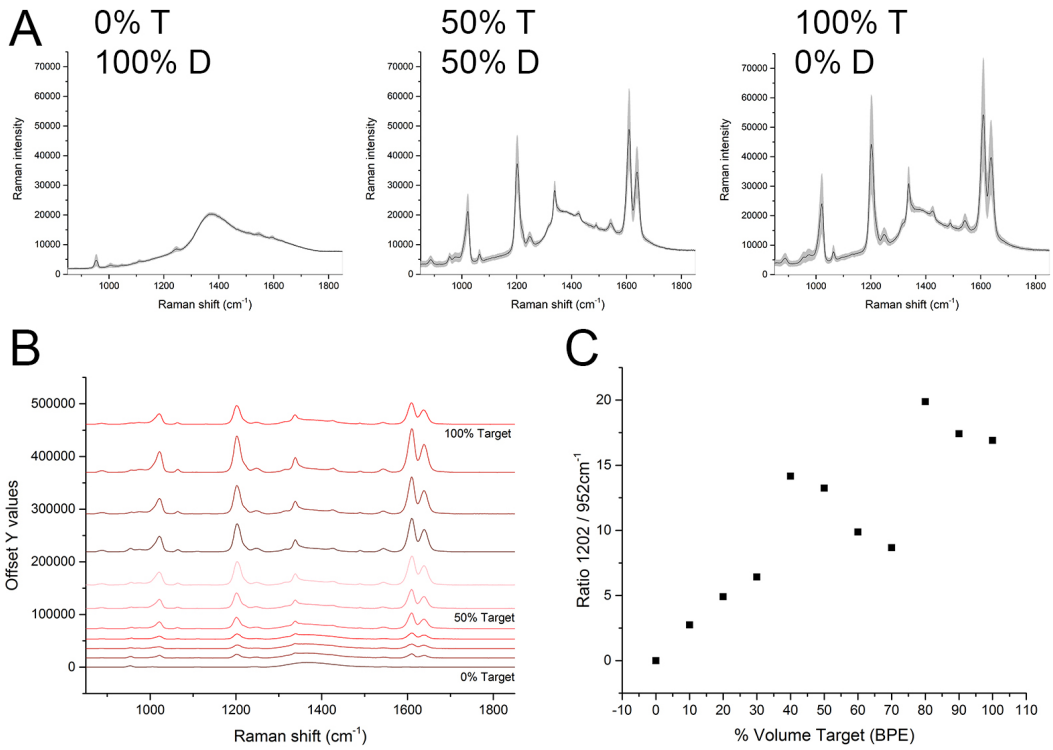


FIGURE 5.11: Reference spectra and corresponding ratios for target and dummy conjugates. A) Raw spectral responses of dried conjugates at the percentage volumes of target (T) and dummy (D) conjugates shown, indicating that the pure dummy conjugate reference peak was not intense enough for ratiometric analysis methods. B) Baselined, averaged spectral responses of dried conjugates at 10% volume intervals. C) Ratiometric reference values for the two primary peaks, 1202cm^{-1} (BPE target) and 952cm^{-1} (PPY dummy).

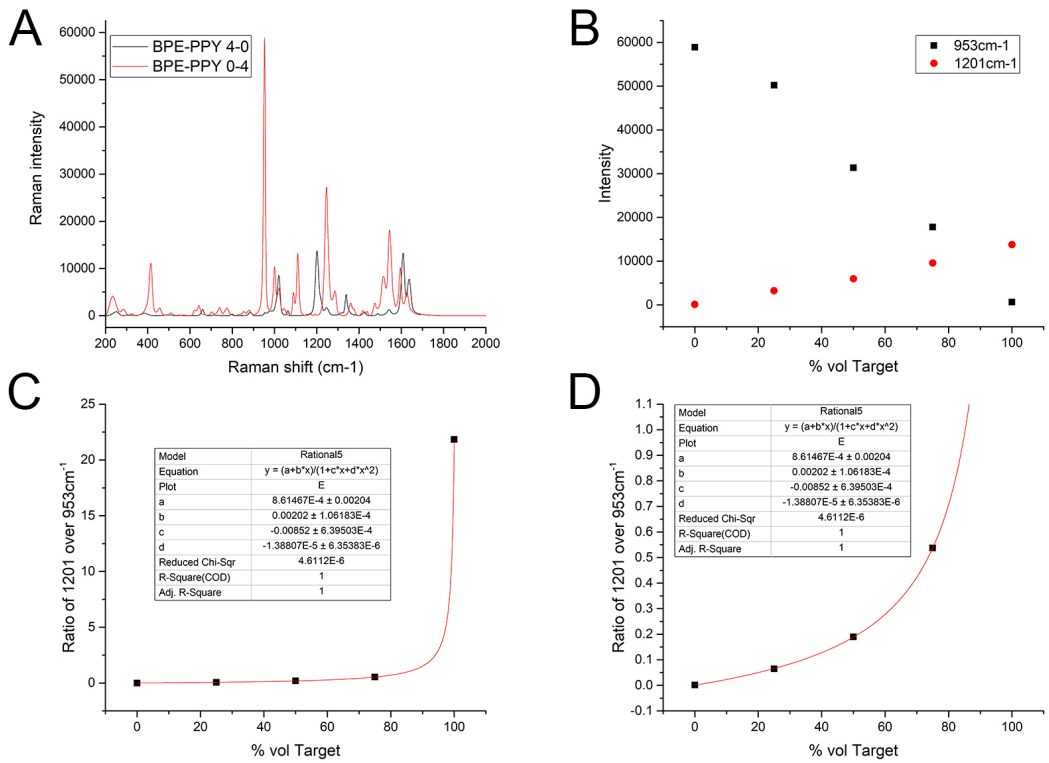


FIGURE 5.12: Ratiometric spectral reference data capable of use with false-colour maps. A) Pure baselined reference spectra of target (BPE-PPY 4-0) and dummy (BPE-PPY 0-4) conjugates, acquired from bulk volumes. B) Intensity values of the reference peaks at 25% interval mixed volumes of dummy and target conjugates. C) Ratiometric value against percentage volume and the corresponding reciprocal fit. D) Graph C on a y-limited scale to account for the near-asymptote of pure target spectra in a ratiometric format, which serves as calibration data for false-colour maps.

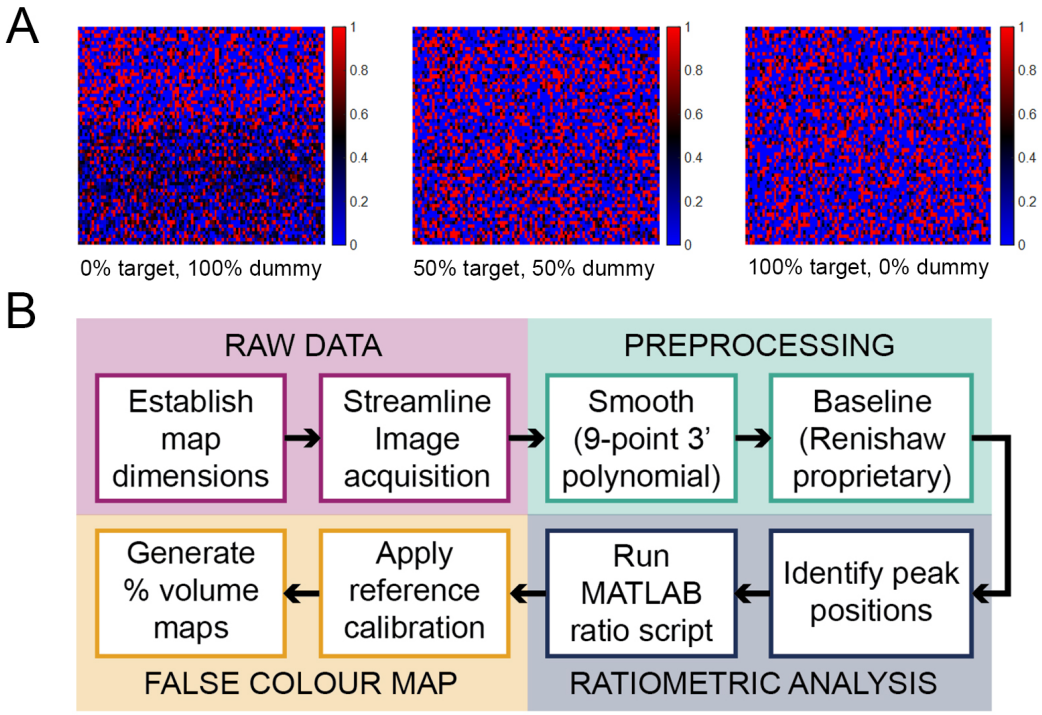


FIGURE 5.13: False colour ratiometric maps, generated using reference peaks 952cm^{-1} (dummy) and 1202cm^{-1} (target). A) Maps of the dried conjugate reference mixtures described above, coloured according to peak ratio. B) Work flow of the signal processing required to perform ratiometric analysis in this context.

5.6 Optimising fluorescent labelling as an analogue for alkyne-tagged antibodies

Following the inactivity of Raman labelled colloidal gold with stimulated Raman scattering, direct antibody labelling with Raman molecules was explored as a vehicle for hybrid stimulated Raman imaging. Investigating a small number of bands specific to those peaks found to be significant in unlabelled cell classification, alongside those associated with the antibody label in question, should allow for a Raman mapping approach which rapidly images robust markers for the disease in addition to global information and morphology for borderline samples.

The Raman-active molecule to be conjugated was N-4-(4-phenyl-1,3-butadiyn-1-yl)benzyl ('BADY')-succinimide (Figure 5.14), synthesised by Dr Will Tipping in Professor Alison Hulme's research group in the Department of Chemistry (UoE). The Raman molecule has been used previously when conjugated to anisomycin for SRS imaging of drug kinetics and phosphorylation of JNK1/2 in breast cancer cell SKBR3 [212]. To optimise the succinimide conjugation protocol, a fluorescent label with succinimidyl ester was first conjugated to an unlabelled secondary antibody so that it could be compared directly to a commercial equivalent.

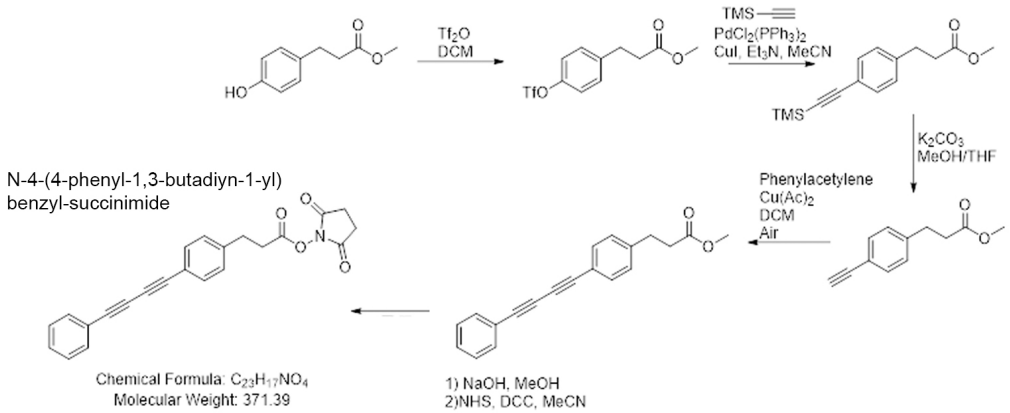


FIGURE 5.14: Synthesis and structure of BADY-succinimide (bottom left), a Raman-active molecule with vibrational peaks in the cellular silent region.

5.6.1 Optimisation of NHS coupling using commercial labels

Preliminary investigations using a 10ml column resulted in a large number of excessively dilute fractions which were not suitable for use with immunofluorescence (Figure 5.15). This was replaced with an 800µl column (35µm pore size, Mobicol #M1002). The antibody was buffer-exchanged and concentrated as described in Section 5.3.2. A 30µl volume of concentrated antibody was mixed with 1.50, 2.25 or 3.00µl fluorescent dye (10mg/ml as described in Section 5.3.2). Difficulty was experienced accurately weighing small masses of unconjugated dye. The absorbance profiles of each sample were collected using 1µl of each sample on a spectrophotometer (NanoDrop 2000c Spectrophotometer, Thermo Scientific). The results are in Figure 5.16 with corresponding degree of labelling (DOL) calculations in Table 5.4. The DOL values were calculated using equations 5.2 and 5.3 in Section 5.3.2. The molecular weight of the antibody is 144g/mol and the extinction coefficient of the dye is $\varepsilon/1000 = 71$.

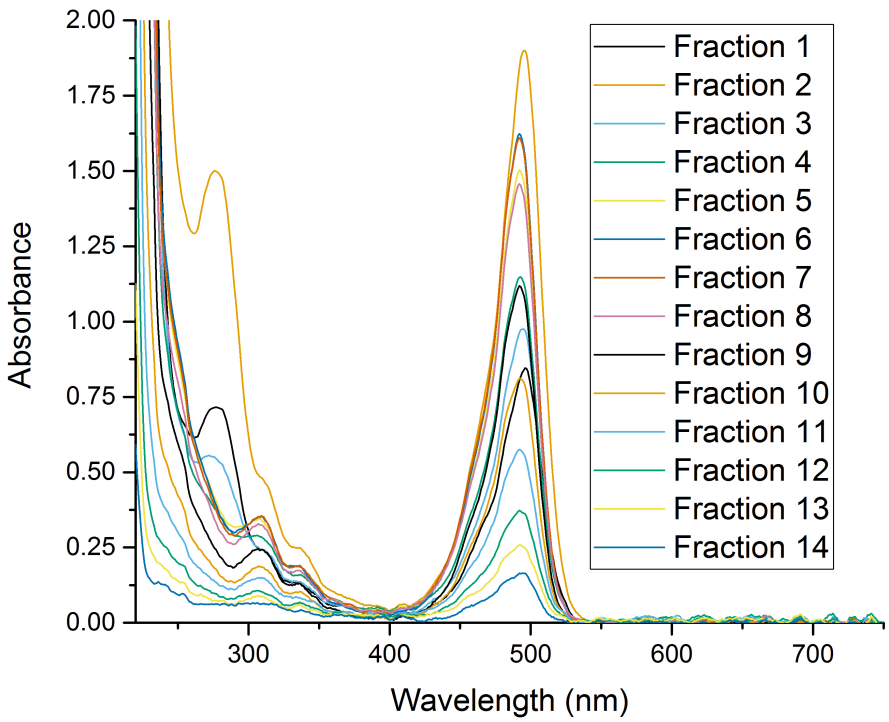


FIGURE 5.15: Absorbance spectra of fluorophore conjugation fractions with 10ml column, which required multiple flushes with PBS to extract the full reaction volume. Fractions 1 and 2 show the highest antibody concentration as expected.

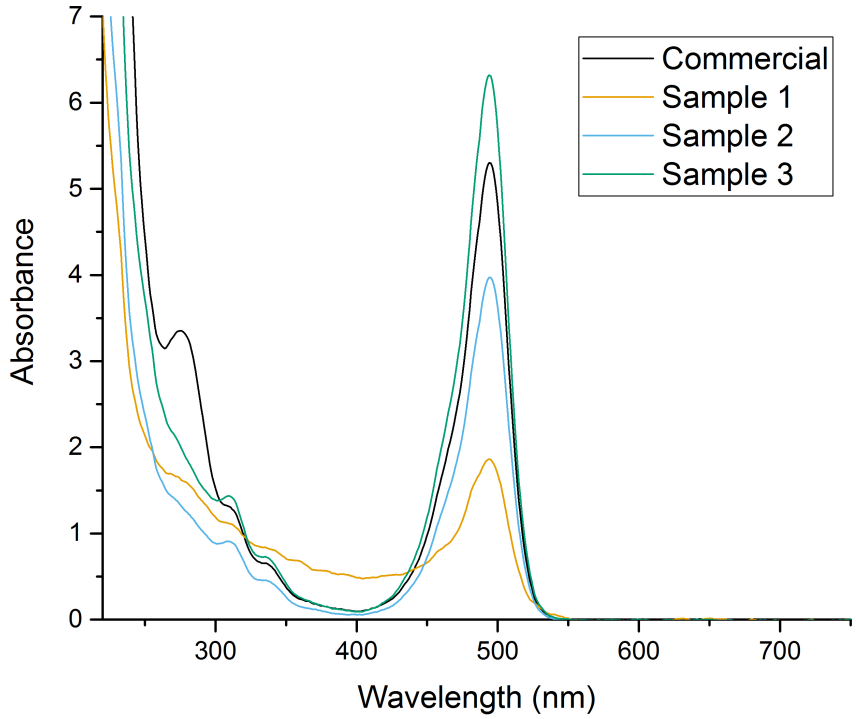


FIGURE 5.16: Absorbance spectra of the three conjugation optimisation samples produced and used for storage experiments, with commercial Alexa Fluor 488 absorbance for reference. Sample 3 shows an excess of dye (absorbance maximum) in relation to protein concentration (280nm).

Sample	A_{496}	A_{275}	A_{mAb}	**[protein]**	**DOL**
Commercial	5.302	3.353	2.769	1.978	5.4
1	1.863	1.619	1.414	1.010	3.7
2	3.972	1.315	0.878	0.627	12.8
3	6.321	1.986	1.290	0.922	13.9

TABLE 5.4: DOL calculations for the preliminary fluorophore conjugation samples produced, based on the absorbance at 496 and 275nm, and protein contribution and concentration.

5.6.2 Lifetime and storage considerations

One of the rationale quoted for the use of SERS as opposed to fluorescence imaging is high photostability over time [167]. The conjugated fluorophores were tested for storage-based signal deterioration on this basis to ensure that SERS offers advantages beyond multiplexed or hybrid imaging in a cervical cancer context.

Wildtype SiHa cells were prepared for immunofluorescence (IF) using the labelled secondary antibody samples produced in Section 5.6.1. The standard IF protocol is described in Section 2.4 - in all cases, BD Pharmingen mouse anti-human antibody to p16 was used as the primary antibody (1:250 dilution) and the produced goat anti-mouse samples used as secondary under the same conditions. The cell slides were imaged 1 day after production and stored at 4°C in darkness for one month before re-imaging. In both cases, the channel exposure times were 23ms for DAPI and 53ms for FITC. The FITC fluorescence signal and population expression for both imaging times can be seen in Figure 5.17.

For non-saturated samples, there is a dip in the normal distribution of fluorescence expression after a month of storage, indicating that probing the intrinsic molecular vibrations of Raman-active tags may provide benefits for clinical samples which require revisiting after storage. Sample 3 demonstrates the opposite trend, possibly related to an excess of fluorophore in the sample.

5.6.3 Comparison of commercial and in-house conjugation using immunofluorescence

The conjugation procedure was repeated with a full aliquot of unconjugated dye (1mg) suspended in 100µl DMSO, before half was decanted and diluted 1:1 in DMSO to ensure exact concentrations of 10mg/ml and 5mg/ml respectively. The absorbance spectra and structure of the unconjugated dye can be seen in Figure 5.18. Eight samples were produced as summarised in Table 5.5 with the corresponding DOL calculations. Sample 8 used frozen dye from the previous experiment to test whether the fluorophore remained stable following reconstitution. The results can be seen in Figure 5.19.

The samples were used to prepare wildtype SiHa for IF, as described in 5.6.2. A comparison of the FITC signal from all 8 samples can be seen in Figure 5.20. Samples 1, 2, 5 and 7, being the closest in DOL to the commercial label, were also prepared with a secondary-only control. These results and corresponding fluorescence histograms can be seen on in Figures 5.21 and 5.22 respectively. Samples 2 and 5 provide the closest DOL to the commercial sample, but the excess of labelling in Sample 2 also generates a

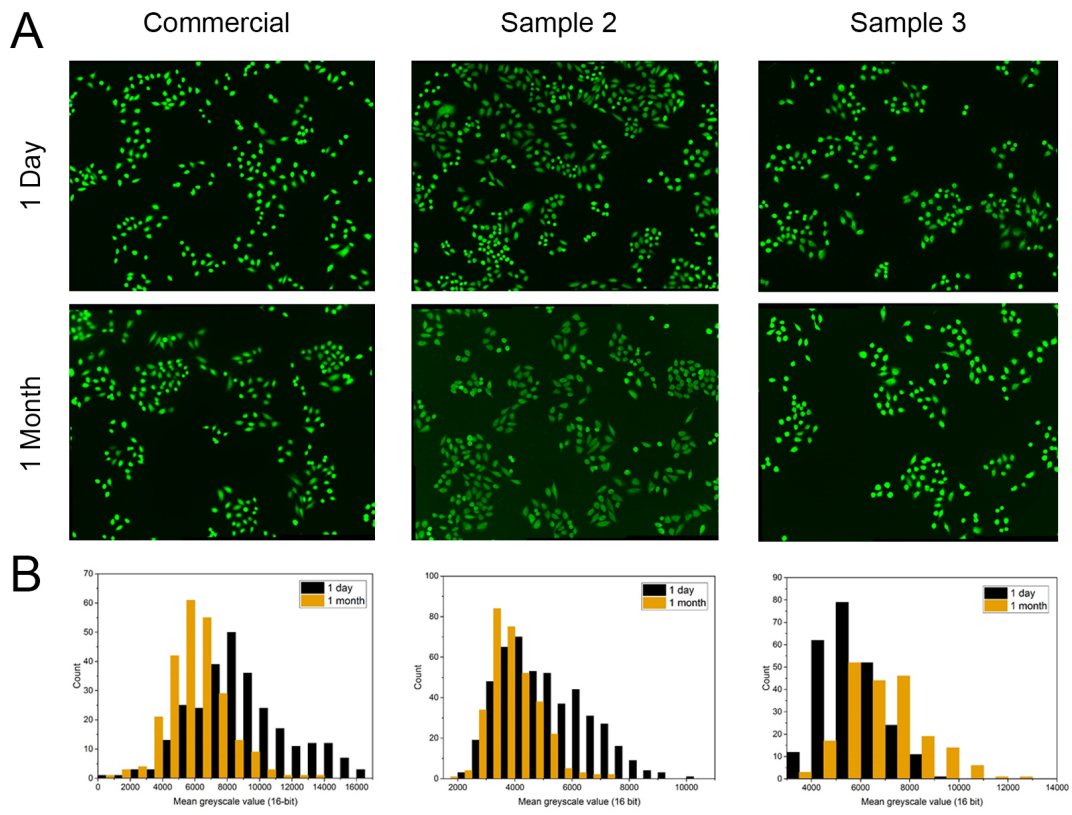


FIGURE 5.17: Immunofluorescence imaging of SiHa w/t cells prepared with commercial or in-house Alexa Fluor 488-conjugated goat anti-mouse secondary antibody. A) Fluorescence maps (FITC channel) of cell expression acquired 1 day and 1 month after preparation. B) Expression histograms of the above data taken from greyscale values of the FITC channel with a nuclear mask, showing that the non-saturated samples (commercial, Sample 2) undergo signal degradation with time.

strong background signal which renders the labelled antibody useless. As such, Sample 5 provides the best indication of labelling ratio for the antibody and succinimide ester, which can be used as a starting point for optimisation of the antibody coupling to BADY-succinimide.

Owing to time pressures, this body of work was passed on to collaborators in the Hulme group so that efforts could be focused on the development of a low-cost, clinically relevant tool.

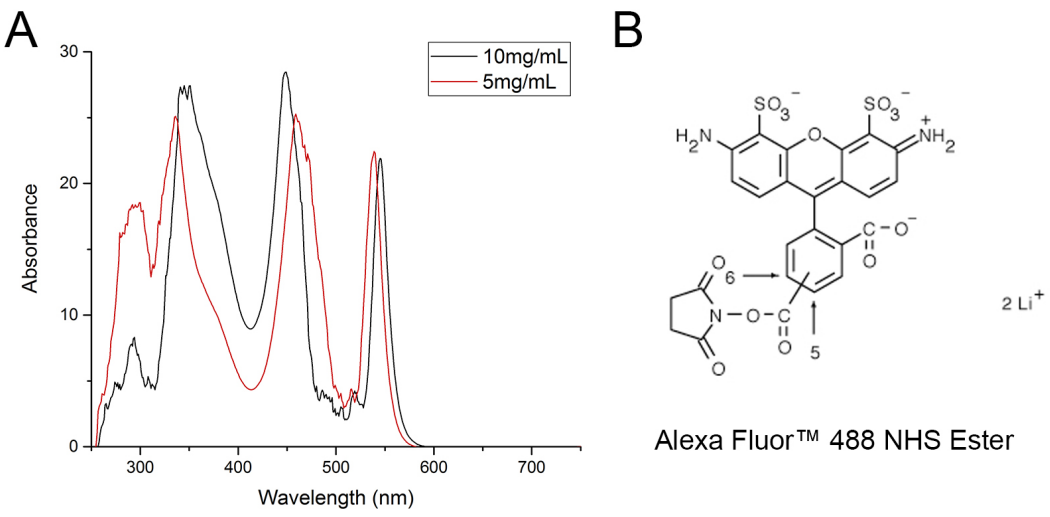


FIGURE 5.18: A) Absorbance profile of unconjugated Alexa Fluor 488 in DMSO at the two concentrations used. B) The structure of the unconjugated succinimidyl dye.

Sample	**Dye (µl)**	A_{496}	A_{275}	A_{mAb}	**[protein]**	**DOL**
Commercial	-	1.881	1.318	1.111	0.794	4.8
1	1.50	1.324	0.678	0.532	0.380	7.1
2	2.25	1.876	1.113†	0.906	0.648	5.9
3	3.00	3.332	1.318	0.951	0.680	9.9
4	1.95	1.719	1.566	1.377	0.983	3.5
5	1.50	0.890	0.776	0.678	0.484	3.7
6	2.25	0.665	1.023	0.950	0.678	2.0
7	3.00	1.668	0.862	0.679	0.485	7.0
8	1.95	6.645	1.471	0.740	0.529	25.5

TABLE 5.5: DOL calculations for the optimised fluorophore conjugation samples produced, based on the absorbance at 496 and 275nm, and protein contribution and concentration. Samples 1-4 used 10mg/ml dye and 5-7 used 5mg/ml. Sample 8 used frozen dye from the previous experiment. † No strong absorbance peak seen.

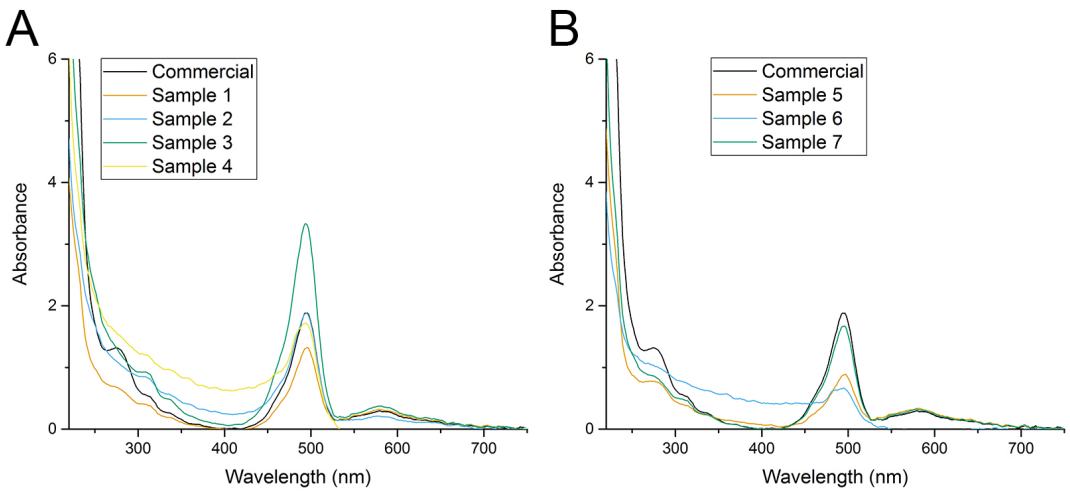


FIGURE 5.19: A) Absorbance profile of the four fluorescent conjugates produced using 10mg/ml dye. B) Absorbance profile of the four fluorescent conjugates produced using 5mg/ml dye. See Table 5.5 for details.

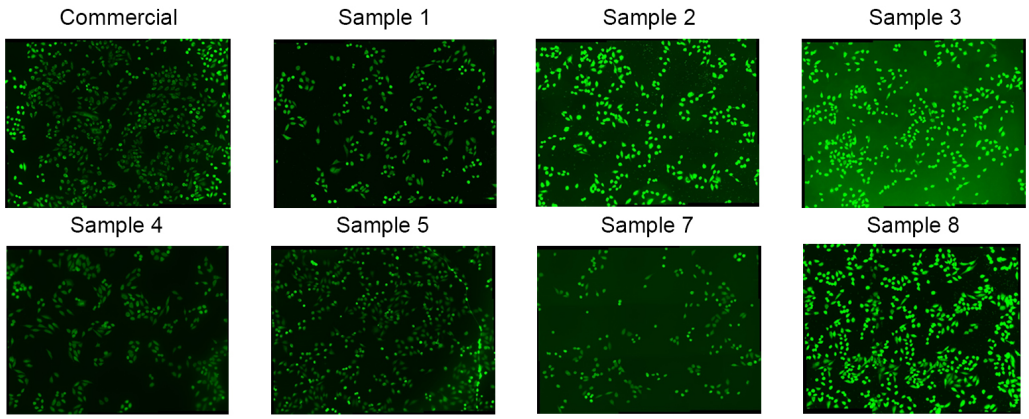


FIGURE 5.20: FITC intensity of SiHa w/t using in-house Alexa Fluor 488-conjugated goat secondary antibodies.

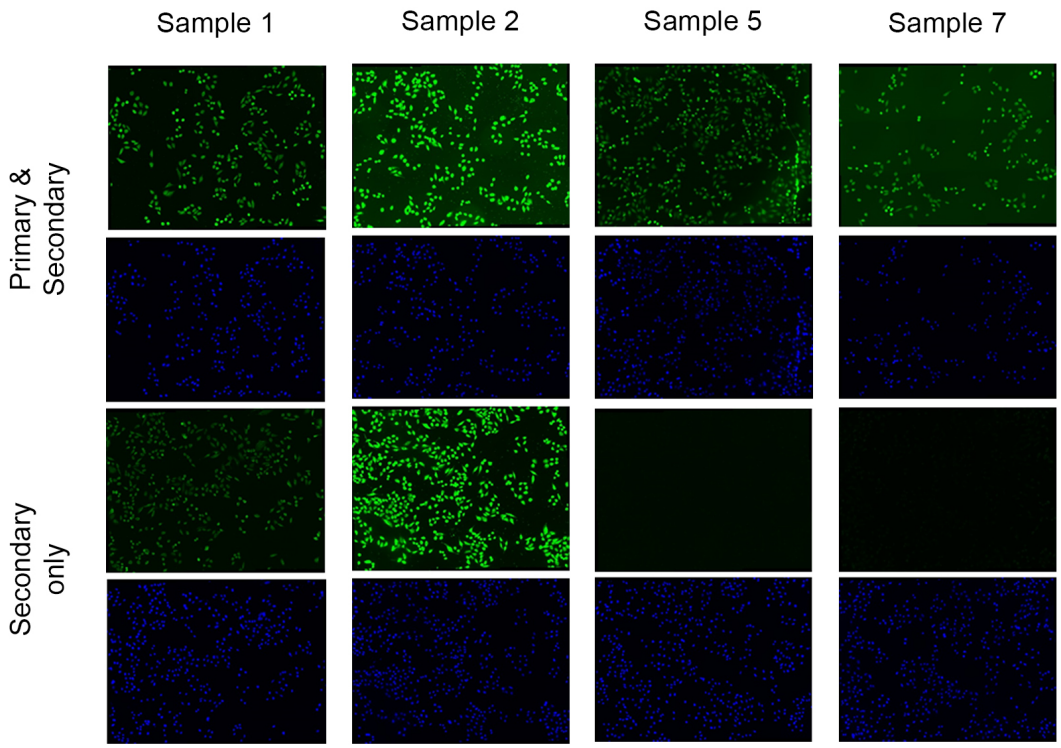


FIGURE 5.21: Optimal in-house Alexa Fluor 488-conjugated goat secondary antibodies compared with secondary-only control. Each FITC image (green) is followed beneath by the accompanying DAPI channel to demonstrate cell position.

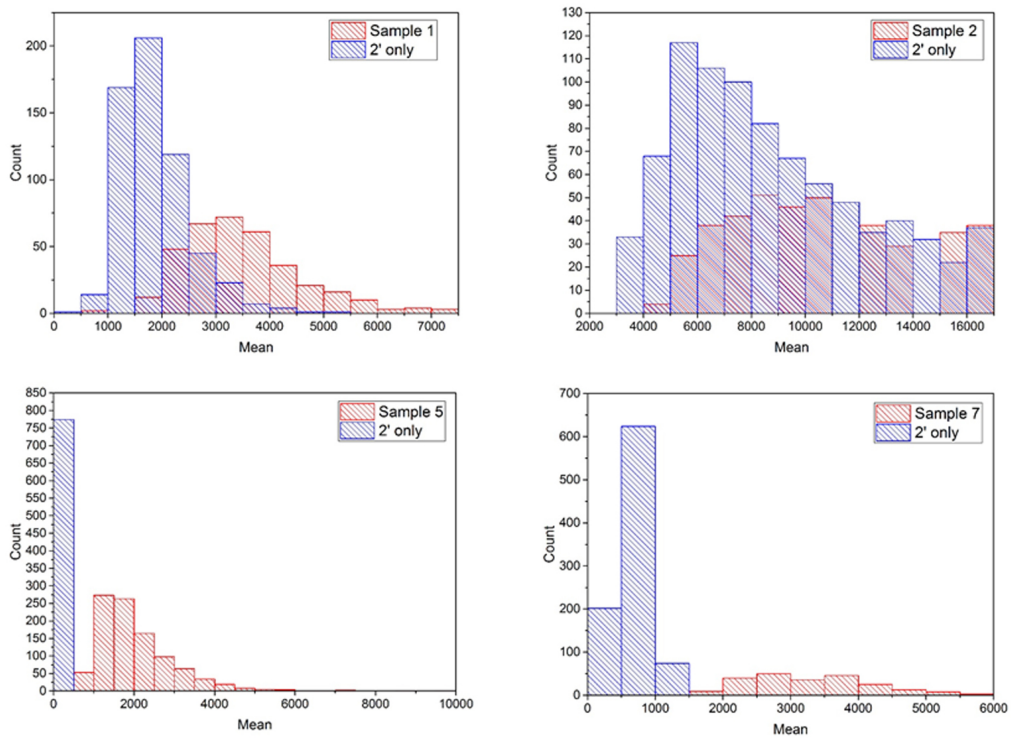


FIGURE 5.22: Fluorescence histograms for optimal in-house Alexa Fluor 488-conjugated goat secondary antibodies, as extracted from fluorescence microscopy images shown in Figure 5.21. This clearly demonstrates Sample 5 as the best-performing conjugate in terms of expression sensitivity and low non-specific binding capacity.

5.7 Conclusions

In this Chapter, three approaches to targeted Raman spectroscopy have been explored as a means of quantifying p16 expression in whole fixed cells.

In the typical SERS approach, a non-specific mimic or 'dummy' conjugate was developed as a complement to the p16-specific target conjugated developed in Chapter 4. Although some issues were experienced with the stability and reproducibility of the dummy conjugate, in-cell experiments demonstrated a clear increase in target signal between wildtype and transfected p16 expression, combined with uncorrelated, small variations in dummy signal as expected. The preliminary true-false analysis of nanoparticle incubation indicated that the dehydrating methanol-based fixative typically used for smear sample preparation provides higher classification accuracy than cross-linking PFA.

Well-fitted spectral reference data has been produced, allowing for the Raman signal from mapping experiments to be calibrated in terms of conjugate ratio. The established signal processing workflow is ready for application with in-cell mapping to generate data which can be quantified with the same confidence as fluorescence microscopy data, rather than by presence of target signal alone.

Owing to the limited scale on which conventional SERS mapping can be applied, SRS-based techniques were also briefly explored as detailed in Appendix B (Chapter 9). Their rapid mapping capability, combined with additional global spectral information, would aid image-guided analysis of cytopathology slides through clearer definition of cell membranes.

The conjugates produced, though biologically stable, were not capable of supporting surface plasmons with pulsed lasers, and so lose their SERS enhancement in an SRS setting. As a result, Raman-active antibody labels were considered. Conjugation parameters were successfully established using a fluorescent analogue bearing the same active group. The method is now ready for development with a synthesised Raman label.

Overall, the global-targeted hybrid imaging approach, employing a strongly Raman-active label in the cellular silent region but with the considerable additional benefit of label-free imaging of cellular structures, offers the most promising approach from a molecular pathology perspective. However, the associated costs of an SRS system would potentially limit its use outside of research and development contexts, at least for the foreseeable future. The following Chapter takes a different approach.

Chapter 6

Development of a sensitive paper-based SERS diagnostic device

6.1 Introduction

Exploratory work performed in Chapter 5 demonstrated overall that while nanoparticle conjugates are capable of indicating levels of p16 expression in individual whole cells, this expression is difficult to quantify in real terms. Although SERS-based ratiometric mapping offers imaging speeds at least an order of magnitude faster than for global Raman interrogation and therefore provides additional spatial data, it lacks both the spectral richness of global spectroscopy and the acquisition scale of other protein-specific imaging techniques such as fluorescence microscopy. Overall, in-cell SERS imaging is advantageous for either highly multiplexed *ex vivo* analysis due to the narrow Raman peaks, or *in vivo* tumour discovery owing to the immune 'stealth' potential of AuNPs compared to other labelling systems [213, 214, 215, 216]. From a molecular pathology perspective cervical smear analysis requires neither of these properties, but rather a test capable of classifying the large and heterogeneous cell population of a single smear according to its progression status.

Thus, a shift in imaging paradigm is proposed towards approaches capable of providing a rapid diagnostic snapshot of whole samples through quantification of a robust marker. As outlined in Chapter 1, this would allow for samples to be rapidly triaged according to p16 expression, so that the volume of samples requiring the diagnostic expertise of a cytopathologist would be reduced to borderline results and routine quality control.

Paper-based biomarker assays or immunochromatographs (ICs) are a well-established, inexpensive and scalable method of clinical whole-sample assessment, the most recognisable example of which is the home pregnancy test [217]. They are primarily based

on two technological principles: the use of paper as a diagnostic platform and the phenomenon of antibody-antigen reactions in a quantitative 'assay' format. The first was explored by Martin and Synge [218] and the second observed by Yalow and Burson [219], for which Nobel Prizes were awarded in 1952 and 1977 respectively. The expiry of the associated patents and their additions in recent years was predicted to generate increased interest in paper assay development for global healthcare endpoints [220]. This is indeed reflected in current academic literature (Figure 6.1) as more research addresses global healthcare issues requiring point-of-care (PoC) solutions [221].

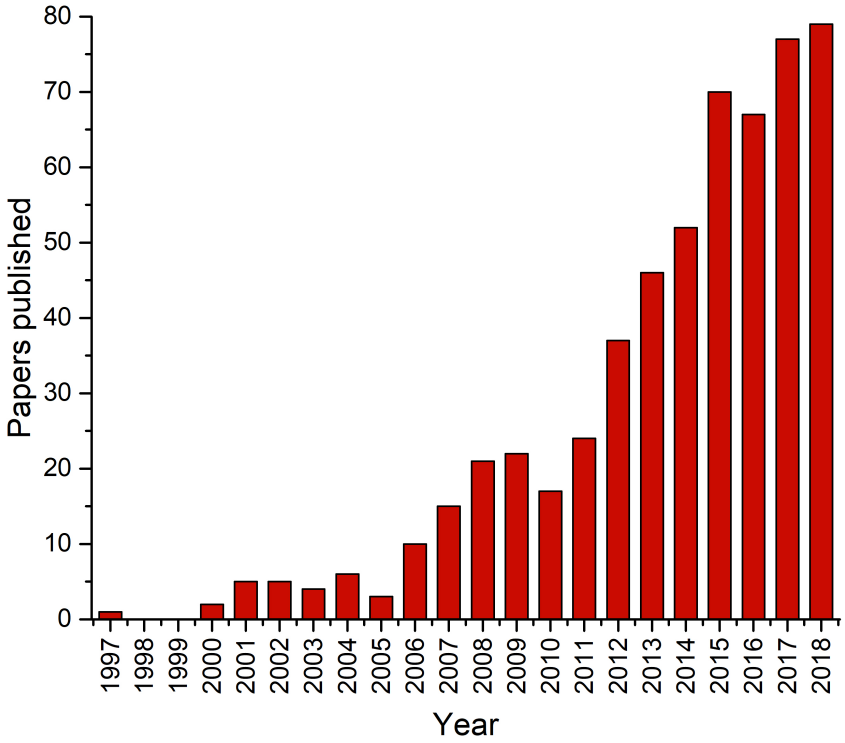


FIGURE 6.1: The number of papers published 1997 to 2018 with key words "lateral flow" and "immunochromatograph*" as generated by Scopus. The lateral flow patent was published in 1997; this publication and its add-ons expired 2006-2013 [220].

ICs have been produced to address diagnostic challenges in situations as diverse as viral, bacterial and fungal infections [222, 223, 224]; food contamination [225, 226]; drug abuse [227, 228], genetic profiling [229, 230] and allergen response [231, 232]. The majority of assays developed in the cervical cancer context are for detection of HPV [233, 234] or proteins thereof [235, 236, 237], possibly reflecting the emphasis on this

high-specificity triage mechanism being implemented in clinical settings. However, detection of p16 has been considered from a paper-based assay perspective. A HRP-conjugated in-house antibody to p16 was passively conjugated to AuNPs, used to classify clinical samples with 85.2% accuracy [187]. An electrochemical sensor for p16 has also been produced, using a complementary pair of unspecified antibodies cloned in-house [186].

Despite this diversity of application the basic layout varies minimally, consisting of a paper substrate including wicking pad, capture lines and a conjugate reservoir (Figure 6.2) alongside storage considerations. Each of these aspects is considered briefly from a technical perspective below. 'Dipstick' and 'lateral flow' assays (LFAs) differ in approaches to conjugate and sample mixing: in the former (Figure 6.2A), the assay is dipped into a well containing a mixture of conjugate and analyte. In the latter (Figure 6.2B), the conjugate is embedded in a dried format in the assay pores, rehydrating as the liquid analyte is added. In printed assays, the assay components are stacked vertically, allowing for multiple tests across the whole 'plate'.

A major additional benefit of paper-based devices, and particularly the LFA, is that these assays possess the characteristics required for PoC diagnostics in low-resource settings. As described in the Introduction, most cervical cancer cases occur in areas associated with limited income and medical infrastructure. Typically, technologies introduced in this setting must be ASSURED: affordable, sensitive, specific, user-friendly, rapid and robust, equipment-free, and delivered to those who need it [238]. Paper-based devices typically adhere to these criteria, providing that they are easily disposed and do not require constant refrigeration.

Two significant challenges in paper-based assay development are flow control and detection sensitivity [239]. The former is a particular issue for an emerging field of more complex designs which fall under 'paper-based microfluidics'. As with powered devices, paper-based microfluidic assays address the need for multiple stages of reagent addition and mixing not required for the proposed use with cervical smears. These are reviewed extensively elsewhere [240, 241, 242, 243, 244].

In the past, detection sensitivity using the intrinsic colorimetric response of nanoparticle based systems has been a limitation when compared to radio immuno (RIA) or enzyme linked immunosorbent (ELISA) assays [245]. It is proposed that the SERS activity of the developed conjugate would allow for a p16 expression assay to be developed that could be used for rapid colorimetric assessment and subsequent triage of samples, as well as a sensitive SERS-based quantification of p16 expression. In the past two decades, SERS PoC assays have not flourished to the same extent as other

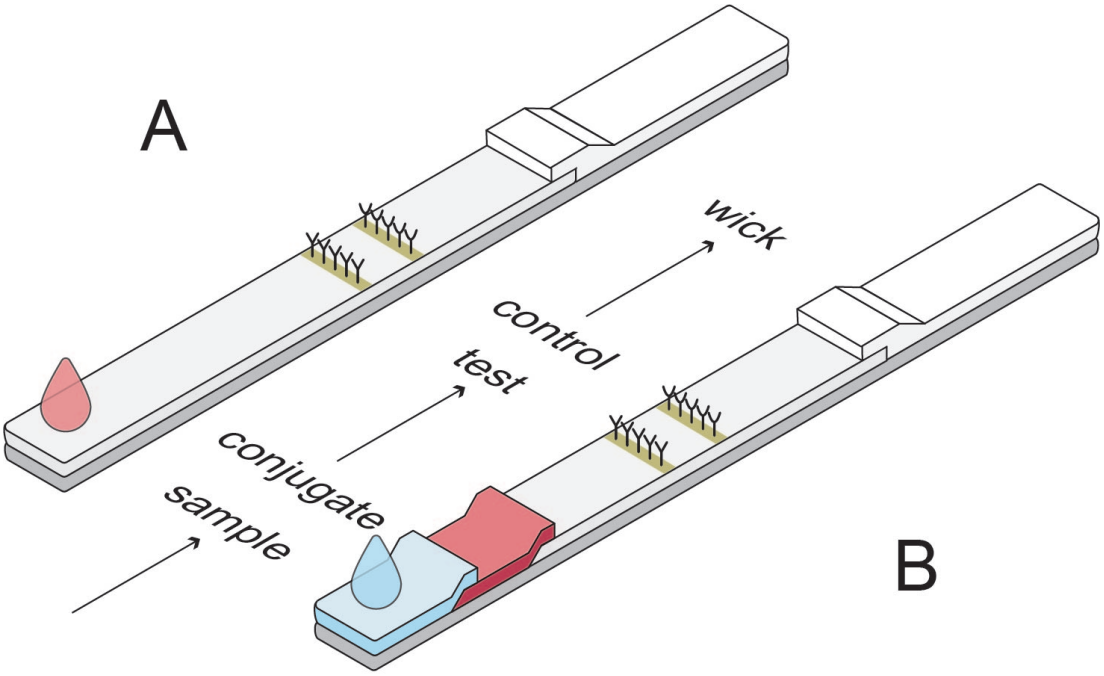


FIGURE 6.2: Schematic of a typical paper-based assay layout. In both cases, the sample and conjugate travel up the strip via capillary flow where they interact with the test and control lines. A) A dipstick assay, where the conjugate and sample are mixed separately before being introduced to the substrate. B) A lateral flow assay, where the sample is introduced to a sample pad before flowing into the dried conjugate pad. This reconstitutes the conjugates and allows the two reagents to react before reaching the test and control lines, as before.

more complex diagnostic assays using this optical phenomenon [246]. However recently, paper-based assays with SERS functionality have been developed for bacterial infection [247], stroke [248] and heart disease [249]. A portable SERS LFA reader has also been developed [250]. Thus far, no successful attempts have been made in the cervical cancer context. A brief summary of the technical considerations of such an assay follows below.

Substrate

The substrate consists of a porous, analyte-binding polymer membrane with a polyester backing for improved handling. Nitrocellulose and nylon membranes employ electrostatic binding. The binding capacity of the membrane is determined by its available surface area, including that of the internal pores, and by the loading capacity of the

protein or other analyte. For nitrocellulose membrane, the peptide bonds of the protein bind to the nitrate esters through strong dipole interactions. The pore size of the membrane determines both the strength of capillary action and the porosity or 'bed volume' of the membrane, which signifies the amount of liquid required to completely wet the device. Overly large pore sizes can cause capture-line spreading and therefore decreased assay sensitivity. Insufficient volumes can result in disruption of capillary flow, resulting in a failed test. Excess sample volumes cannot be incorporated into the device structure and so will reduce the overall signal.

Capture lines

For both colorimetric and SERS-based quantification, the capture lines should be printed at the membrane surface to prevent loss of signal to membrane depth. Proteins bind almost instantaneously on contact with nitrocellulose, and so the wetting properties of the streptavidin binding layer and capture peptides should be controlled such that narrow lines are produced. A narrower capture region results in a more concentrated and therefore optically sensitive signal.

The use of streptavidin anchoring on nitrocellulose membrane has been reported as insufficiently robust for lateral flow assays, and so use of nitrocellulose-binding mutant streptavidin may be required [251].

Sample flow and detection

Owing to the often homogeneous nature of the membrane's pore structure and the exponential decay of capillary flow rate with distance travelled, substrates are characterised by their capillary flow time - the time taken for the liquid sample of interest to completely fill a strip of defined length. Longer flow times typically result in higher sensitivities, as the ligands have more time to interact with their target.

The addition of a sample pad to membranes can fulfil several functions, from simply allowing a more controlled distribution of sample to adding other reagents. In the case of a lateral flow assay, the sample liquid must then hydrate and interact with nanoparticle conjugates, stored in the conjugate pad. The conjugate pad itself must not bind to the nanoparticle complex. Slow release of conjugates is preferential, as a colorimetric signal will only be produced by analyte flowing with or ahead of the conjugates.

If a paper-based device for p16 quantification were to incorporate SERS-active conjugate, it could easily be combined with a handheld Raman spectrometer with 'point

and shoot' attachment. This would allow for rapid, quantifiable triage of smear samples in low-resource settings without the requirement for a trained cytopathologist. Raman interrogation requires optimisation of laser power and acquisition times and cycles to ensure a rapid test which does not burn the membrane.

Storage

Storage is a key consideration for paper-based devices - the completed device must be suitably robust to withstand transportation and handling within ASSURED criteria whilst maintaining biological functionality of the peptide and antibody components. This is the case for both temperature and drying considerations. Antibody-conjugated nanoparticles and capture layers must maintain their structural integrity when dried, and be capable of binding immediately upon rehydration with the sample complex. Drying is often carried out in the presence of sucrose or trehalose, which forms a protective layer around the protein molecules [252].

6.2 Concept: paper-based diagnostic devices

As described above, a typical dipstick or lateral flow strip has four key physical features; the paper substrate, wicking pad, conjugate reservoir and capture regions, which consist of a test and control line.

The test line must be populated by a ligand capable of binding to the target molecule of interest - in this case, p16 protein - when it is in complex with the nanoparticle-tagged antibody conjugate described in Chapter 4. Binding kinetics must be sufficiently rapid for the target to be captured within the time taken for the sample volume to run across the test line. In addition, the binding ligand must be specific for the target in question; competition would reduce the colorimetric and SERS signal at the test line as the conjugate label is specific for the target. A complementary antibody could fulfil this role; currently none are known to be available.

The control line must simply consist of a ligand capable of binding to the nanoparticle-tagged antibody conjugate when it is not bound to p16, to ensure that negative test results are not merely the result of a failed assay. In commercial conjugate assays this often takes the form of a Protein A/G mix, which binds to the heavy chains of a variety of IgG types from mammalian hosts and is therefore largely universal. However, the epitope of p16 itself also represents an ideal candidate. Provided that the capture

line is printed upstream of the test line to prevent disassociation of control line mimotopes and subsequent false positive results, the use of the p16 epitope would confirm in each test that the antibody conjugate is still biologically functional.

In summary, a successful paper-based device would require ligands capable of binding to p16 antigen and p16 antibody in a rapid and specific manner. The ligand to p16 itself should be specific for a different epitope than that already used. Conversely, the ligand to p16 antibody should mimic the epitope already used for conjugate binding (Figure 6.3).

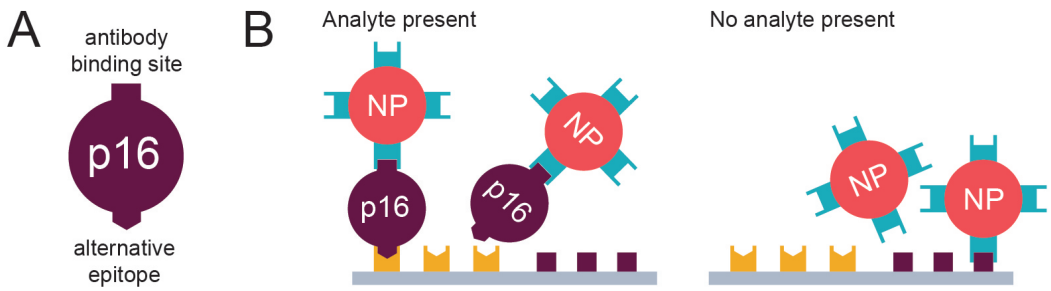


FIGURE 6.3: A schematic of p16 epitopes and ligands in the context of a paper-based assay design. A) The 'antibody' and 'antigen' epitopes of p16. B) Interaction of analyte-bound and unbound nanoparticle (NP) conjugates with the paper-printed test line (yellow) and control line (purple) peptides.

One method capable of elucidating the protein motifs described above is phage-displayed peptide biopanning, referred to as 'phage display' or 'biopanning' for convenience. Developed by George Smith and Sir Gregory Winter in the 1980-90s, it was awarded the 2018 Nobel Prize in Chemistry [253]. Its ability to link peptide affinity to DNA sequence inexpensively in a broad range of biomedical contexts [254] has led to significant advances in biopharmaceuticals in particular, including the first six commercialised antibody medicines [255].

In this method, a large population of viruses or 'phages' capable of infecting bacteria are genetically engineered to include a randomised insert in their genome which produces a short protein sequence or 'peptide' that is displayed on the outside of the bacteriophage at the end of the pIII coat protein. Libraries typically contain 10^9 unique sequences, with many copies of each sequence within a working volume. The random library of displayed peptides can then be panned against targets of interest (Figure 6.4) over multiple rounds to enrich the library for peptides specific to the target.

In the panning process, the target is coated onto a polyester substrate and the phage library introduced for peptide binding to take place. The system is washed

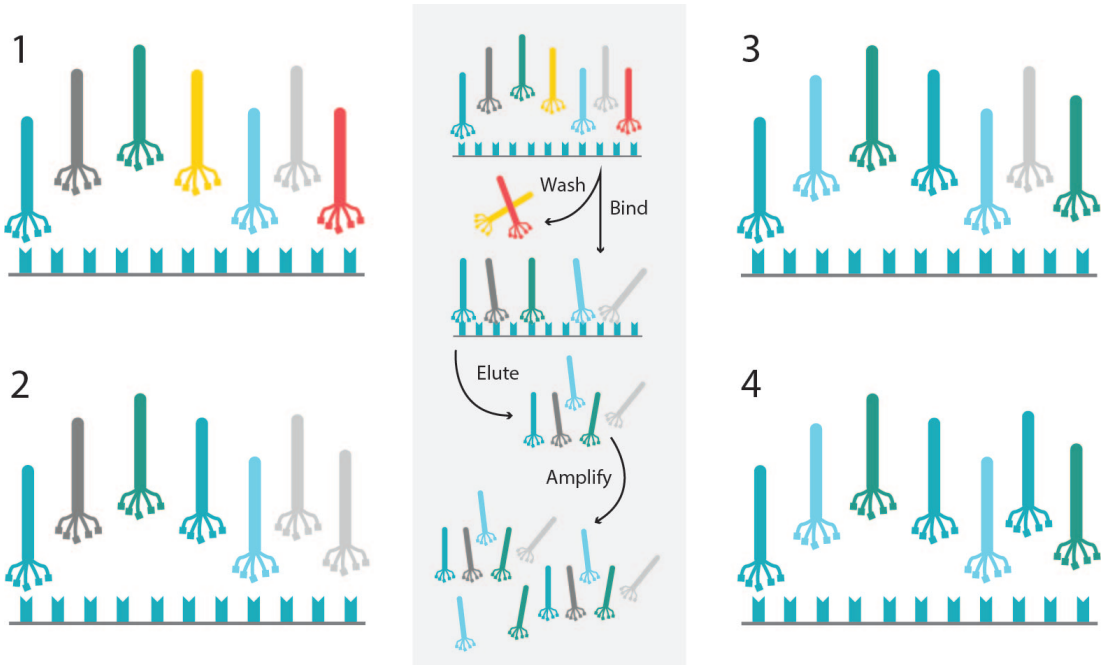


FIGURE 6.4: A schematic of the full biopanning process, including the panning protocol (grey panel) and library enrichment for consensus, high-affinity peptide motifs over Rounds 1, 2, 3 and 4.

to remove unbound phage before the remaining phages are eluted and introduced to the appropriate bacterial culture. The bacteria replicate, amplifying the eluted phage pool. The pool, once purified, can be reused as an enriched library to narrow the population for target 'hits'. By sequencing the altered phage genome, the corresponding DNA and protein sequence can be found. Thus, biopanning represents a convenient method for discovering novel binding ligands.

In the context of this paper-based device, the peptides need to mimic the binding conditions described above. Once the DNA sequence is known, the peptides can be synthesised with a spacer and biotin tag at the C-terminus. The addition of a biotin tag allows the peptides to be oriented consistently during affinity testing by binding to a layer of streptavidin immobilised on the assay plate. Critically, it also allows them to be fixed to a printed line of streptavidin during paper-based assay preparation, reducing the likelihood of disassociation from the membrane during flow and binding.

6.3 Materials and methods

Sterilised tips were used for all procedures and all washes pooled and treated with Virkon to eliminate contamination. Ultrapure water ('dH_2O', MilliQ) was used for all experiments. All plate shaking steps were carried out at room temperature (RT) on a benchtop plate shaker (IKA MS 3) set to 500rpm and plates sealed with Parafilm M for all incubation steps. Centrifugation steps were executed at 4°C. Phosphate buffered saline ('PBS', pH7.3) consists of 137mM sodium chloride (NaCl), 3mM potassium chloride (KCl), 8mM disodium hydrogen phosphate (Na_2HPO_4) and 1.5mM potassium dihydrogen phosphate (KH_2PO_4). For ease of reference, all other reagent components are listed in Table 6.1.

Buffer type	**Supplier (Cat. #)**	**Constituents**	**Storage**
Lysogeny broth	In-house	1.0% tryptone, 0.5% yeast extract, 1.0% NaCl, 0.1% glucose [w/v]	RT
Tetracycline solution	Sigma Aldrich (T7660)	5mg/ml tetracycline hydrochloride in 1:1 dH_2O:EtOH	-20°C, dark
IPTG/Xgal stock	BioVision (1649-1G), Formedium (XGAL001)	1.25g isopropyl-β-D- thiogalactoside, 1.00g 5-Bromo-4-chloro-3-indolyl-β-D-galactoside	-20°C
Sodium bicarbonate buffer	Sigma Aldrich (S5761)	0.1M $NaHCO_3$ (pH8.6)	RT
Tris-buffered saline (TBS)	VWR	50mM Tris-HCl (pH7.5), 150mM NaCl	RT
Blocking buffer	Melford (A1302)	Albumin Bovine Fraction V (BSA) in sodium bicarbonate buffer*	4°C
Iodide buffer	VWR (SS0620)	10 Tris-HCl (pH8.0), 1 EDTA, 4M sodium iodide	4°C, dark
PEG/NaCl	Sigma Aldrich (P5413), VWR	20% PEG 8000 [v/v], 2.5M NaCl	RT, sterile

TABLE 6.1: Constituents of buffers and other reagents used during the peptide discovery process. Tris-HCl is abbreviated nomenclature for tris(hydroxymethyl)aminomethane ($(HOCH_2)_3CNH_2$) hydrochloric acid (HCl). *BSA concentration in blocking buffer varied depending on experiment iteration (see main text for details).

6.3.1 Phage-displayed peptide biopanning

Phage-displayed peptide panning was carried out with a Ph.D.^{TM}-12 Phage Display Peptide Library Kit (New England Biolabs, Cat. # E8110S) and recommended Tetracycline-resistant *E. coli* strain K12 ER2738, with genotype

$$F'proA^{+}B^{+}lacl^{q}\delta(lacZ)M15zzf :: Tn10(Tet^{R})/$$

$$fhuA2glnV\delta(lac-proAB)thi-1\delta(hsdS-mcrB)5$$

The supplied randomised library of 12-mer peptides are expressed at the N-termini of the pIII minor coat proteins of individual M13KE filamentous bacteriophages, amplified to give approximately 10 copies of each insert per 1µl of phage library. The pIII protein binds to the F-pilus of *E. coli*. The M13KE phage vector (Figure 6.5) is derived from M13mp19, which carries the lacZα gene. In the presence of IPTG, a functional β-galactosidase enzyme hydrolises Xgal to form 5,5'-dibromo-4,4'-dichloro-indigo, producing blue plaques in a blue-white screen upon successful infection of ER2738.

PBS and NaCl/PEG mix were autoclaved before use. In all cases, lysogeny broth (LB) was used with a 1:1000 dilution of Tetracycline solution (abbreviated to 'LB/Tet') to prevent bacterial contamination of the ER2738 competent *E.coli* cells with other wild-type ('w/t') lines.

Biopanning

This protocol has been slightly adapted from that provided in the kit manual due to a combination of the experiment requirements and issues with low phage titre following the wash protocol.

A 96-well plate was washed with Tris-buffered saline with 0.1% Tween 20 (0.1% TBS-T), 1µl p16 monoclonal antibody ('mAb', BD Pharmingen, 0.5mg/ml) added in 199µl sodium bicarbonate buffer for each experiment and the plate incubated overnight at 4°C. For the first round of biopanning, mAb solution was removed from one well (antigen panning, 'AN' samples), washed three times with 0.1% TBS-T and 1µl purified p16 added (100ng/µl in PBS) in 199µl sodium carbonate buffer for 2h at room temperature (RT). Antibody was retained in the remaining well for antibody panning, referred to as 'AB' samples. Both wells were washed as above and filled with 250µl blocking buffer with either 30 or 5mg/ml bovine serum albumin (BSA) [w/v] for 30min with shaking before washing three times with blocking buffer (see Sections 6.4.1 and 6.4.2 for details).

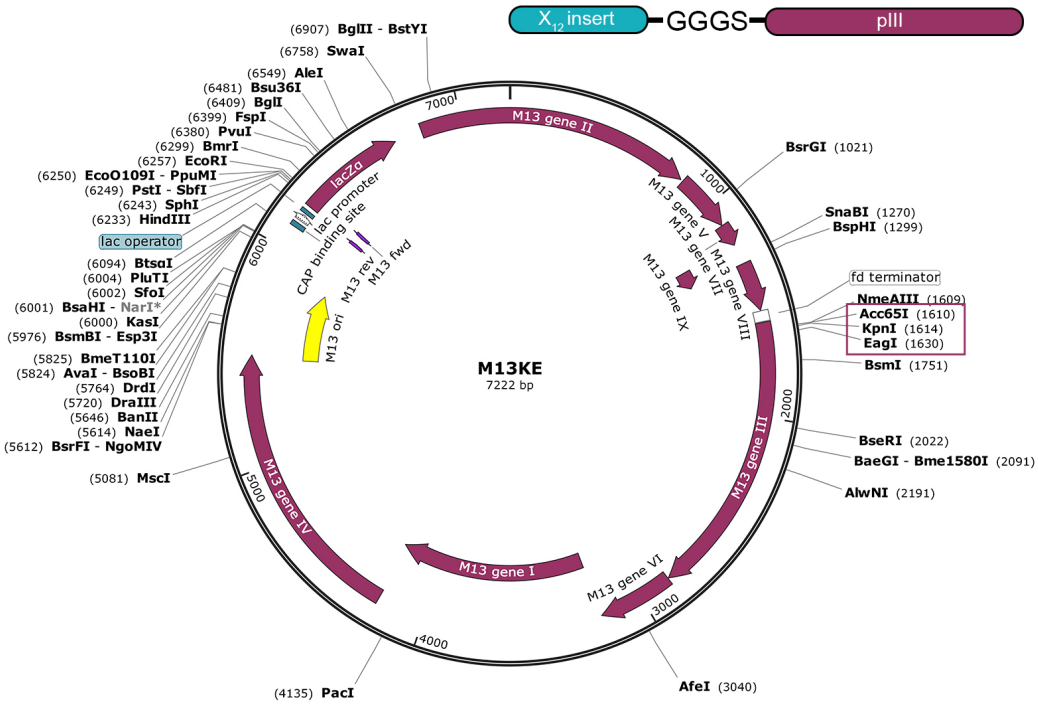


FIGURE 6.5: Sequence map of M13KE bacteriophage, used for the biopanning process. The library insert sequence appears in between Acc65 I/Kpn I and Eag I (purple box), at the end of the pIII minor coat protein, preceded by an additional serine residue and followed by a GGG spacer (top right).

For the first round of panning, the randomised Ph.D.-12 peptide library was introduced to the AB well for 1hr with shaking in order to 'pre-clear' any peptide sequences with high affinity for the BD Pharmingen antibody before adding to the AN well, so as to reduce the likelihood of false hits for p16. This was then transferred to the AN well under the same conditions, with the AB well filled with TBS-T to prevent drying. Both wells were washed five times with blocking buffer before the remaining, bound phage were eluted with shaking for 20min using 200µl 0.1M glycine (pH2.5)+1mg/ml BSA [w/v]. The elutants were transferred to sterile microfuge tubes, neutralised with 20µl 1M Tris (pH8.8) and stored temporarily at 4°C for amplification.

The second round of panning was conducted as described above but using the amplified phage eluate (see below) instead of the randomised library and without the pre-clearing step. Third and fourth rounds followed the same protocol as the second, but with TBS-T washing buffer containing 0.5% rather than 0.1% Tween-20 to remove weakly bound phage.

Phage amplification

Prior to amplification, 10ml LB/Tet medium was inoculated with ER2738 stock (stored at -80°C) in a closed 50ml Falcon tube and incubated overnight at 37°C in a Heraeus Function Line incubator. The resulting stock culture was maintained throughout the panning experiment by diluting every 24hr 1:100 in LB/Tet.

For amplification, the stock culture was diluted 1:100 in 20ml LB/Tet in 125ml Erlenmeyer flasks to ensure adequate aeration. The two unamplified eluates, AB and AN, were added to two separate flasks and incubated at 37°C for 4.5hr at 200rpm. Flasks were loosely coated with foil to reduce likelihood of contamination with wild-type phage. The phage-infected cultures were then transferred to 50ml Falcon tubes and centrifuged at 3184g for 15min to pellet the ER2738 cells. The phage supernatants were transferred to fresh tubes and the process repeated for 2min. The upper 80% of supernatant was transferred again and PEG/NaCl added at 1/6 volume of the supernatant. The solution was allowed to precipitate overnight at 4°C.

The PEG/NaCl precipitations were centrifuged at 3184g for 15min. The supernatant was carefully removed and discarded, and the white phage pellet suspended in 1ml TBS and transferred to a microfuge tube. The samples were centrifuged again at 16100g for 5min to pellet residual cells and the resulting supernatant precipitated in a fresh tube with 1/6 volume PEG/NaCl for 1hr on ice. The precipitations were centrifuged at 16100g for 10 minutes, the supernatant carefully removed and discarded and the phage pellet suspended in 200µl TBS. The amplified eluate from the previous round was used in place of the randomised library for the second, third and fourth rounds of panning.

Eluate titration

IPTG/Xgal plates were produced for blue-white plaque selection and stored at 4°C until use (plates were discarded if contaminated or discoloured). For this, 15g Ultrapure Agarose (Invitrogen 16500500) was dissolved by heating in 1l LB/Tet. Once the LB/Tet/Agar was cooled below 60°C, IPTG/Xgal stock was added at 1:1000 dilution and the mixed solution poured rapidly into bacterial culture plates (Thermo Scientific Sterilin Single Use Plastics) at approximately 20ml per plate. Once cooled and set, the prepared plates were warmed at 37°C in an incubator for at least one hour before use.

To titrate phage samples, 10ml LB/Tet was inoculated with 1:100 dilution of ER2738 stock and the bacterial cells incubated for 4-8 hours with shaking at 200rpm until no

more than mid-log phase was reached ($OD_{600} \leq 0.5$), confirmed by spectrophotometry (NanoDrop 2000c, Thermo Scientific) using 1ml of sample in a cuvette. An overconfluent culture was found to produce 'lawns' of bacteria upon plating, rather than discrete plaques. Overconfluent cultures were diluted with LB until a suitable optical density was reached.

Shortly before plating, 'top agar' was produced as a growth matrix for ER2738 by dissolving 7g UltraPure Agarose, with heating, per 1l LB required. Immediately before use to prevent premature solidification, 3ml top agar was dispensed into 15ml Falcon tubes per plate to be titred. The bacteriophages were serially diluted in 1ml volumes of LB, typically ranging 10^{10} to 10^{20} or 10^{2} to 10^{6} for amplified or unamplified phage pools respectively. 10µl of each serial phage dilution was added to 200µl mid-log phase ER2738, vortexed briefly and left to incubate for 5min at RT before adding to the dispensed volumes of top agar. This preparation was inverted very briefly to mix before being poured onto labelled IPTG/Xgal plates and swirled to distribute. Once set, the plates were inverted and incubated overnight for no more than 18hr to allow for plaques to form.

To calculate the number of plaque-forming units per microlitre (pfu/µl) or more typically per 10µl volume, the plaques on plates with no more than 100 blue plaques were counted and the number multiplied by the inverse of the dilution for that plate. For example, a plate with 34 plaques produced with serial phage dilution of $1{:}10^{9}$, or 10^{-9}, would contain 34×10^{9} plaque-forming units per 10µl sample used, equivalent to 34×10^{8} pfu/µl. Typical examples of titred plaques can be seen in Figure 6.6.

Selection and characterisation

Titred phage plaques were selected for individual amplification following the fourth round of biopanning. Here, 3ml dispensed volumes of 1:100 ER2738 stock in LB/Tet were inoculated with individual well-separated plaques, picked from the plates using sterile pipette tips. The resulting infected ER2738 were left to multiply for 4.5hr at 37°C. The tubes were centrifuged at 3184g for 2min to pellet ER2738 cells. The resulting supernatant for each phage sample was transferred to sterilised Eppendorf tubes: a 1ml sample for DNA extraction, a 1.5ml sample for short-term use including ELISA (stored at 4°C) and a 0.5ml sample for long-term storage, diluted 1:1 with sterile glycerol and stored at -20°C. Affinity ELISA protocols are discussed in Section 6.3.2.

Phage DNA was extracted for sequencing using a variant of the 'Rapid Purification of Sequencing Templates' protocol supplied with the phage display kit. 1ml amplified

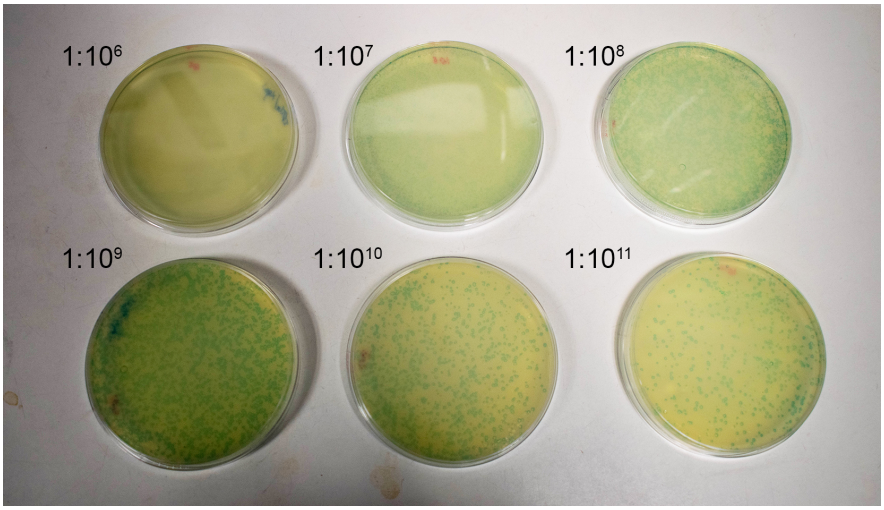


FIGURE 6.6: Typical example of successful blue/white plaque titration. No white plaques are seen, indicating a wildtype contaminant-free pool. Lower dilutions ($1:10^6$) result in a faint blue 'lawn' of plaques, whereas higher dilutions yield well-separated and quantifiable plaques.

phage supernatant per sample was mixed with 400µl PEG/NaCl solution and precipitated on ice for 45min. The precipitate was centrifuged for 10min at 16100g and the supernatant carefully removed. The pellet, often difficult to visualise, was resuspended thoroughly in 100µl iodide buffer and added to 250µl EtOH. The phage DNA was preferentially precipitated by leaving for no more than 20min at RT, before centrifuging at 16100g for 10min. The resulting phage DNA pellet was washed once in ice-cold 70% EtOH [v/v] before re-spinning, discarding the supernatant and drying the pellet at RT. The DNA pellet was suspended in 25µl TE buffer and sequenced in-house on a 3730 DNA Analyzer (Applied Biosystems, Thermo Scientific) using -96gIII reverse primer (5'- CCC TCA TAG TTA GCG TAA CG -3').

The library insert sequence and conserved surrounding sequence of M13KE Gene III can be seen in Figure 6.7. This information was used to identify the library insert sequence and interpret the corresponding amino acids, shown in Figure 6.8.

6.3.2 ELISAs

Enzyme-linked immunosorbent assays (ELISAs) were used to determine the binding affinity of native phage and synthetic biotinylated peptides without the additional complexity of paper-based assay formats, which require significant optimisation in their own right. ELISAs were conducted in Costar 96 Well White Polystyrene High

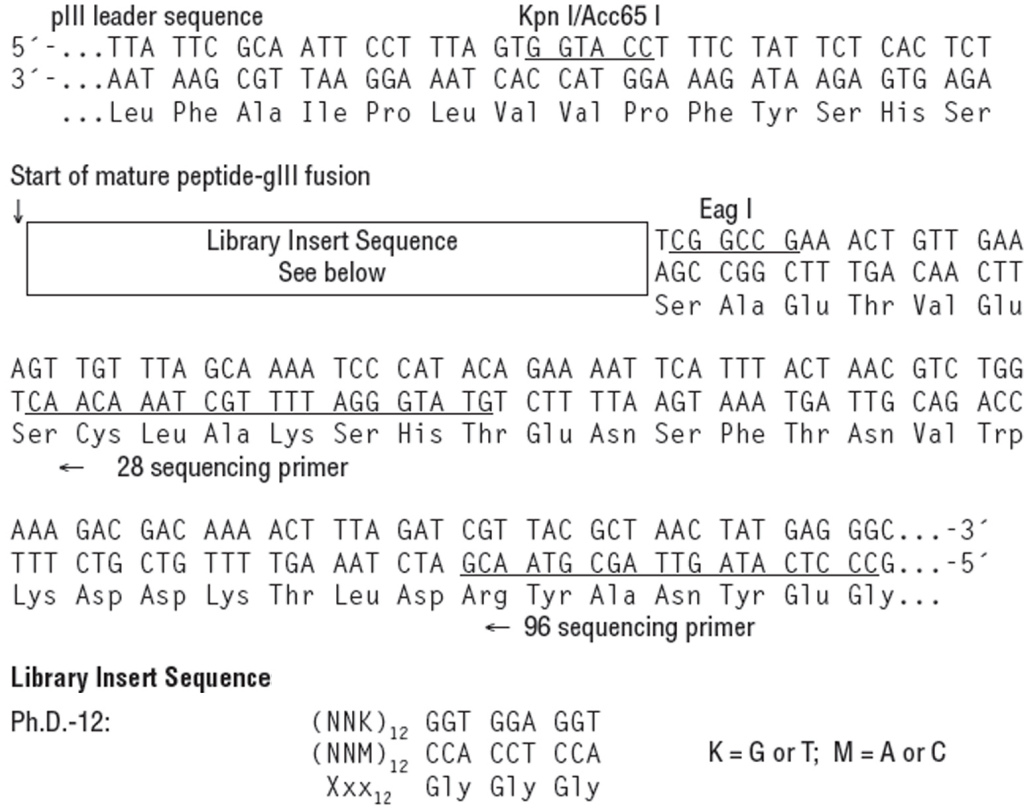


FIGURE 6.7: The phage library insert sequence and surrounding M13KE phage vector. The enzyme cutters Kpn I/Acc56 I and Eag I were used to rapidly identify the position of the insert from sequencing data. The third base in each codon of the insert is always G or T (3′ to 5′).

Binding Assay Plates (#3922). All washing steps were carried out with 150µl 1% Tween 20 in PBS (PBS-T) or TBS (TBS-T) per well, repeated five times unless otherwise stated. The solution was removed through rapidly inverting plates onto paper towels. ECL solution was produced in-house as described in Chapter 2.

Native phage peptide affinity ELISA

Following an initial amplification of individual phage clones as described in Selection and Characterisation, Section 6.3.1, the phage titres were further enriched to test for target specificity as recommended in the kit protocol. A native phage ELISA, where the peptides are displayed on the pIII minor coat protein of the virus rather than being

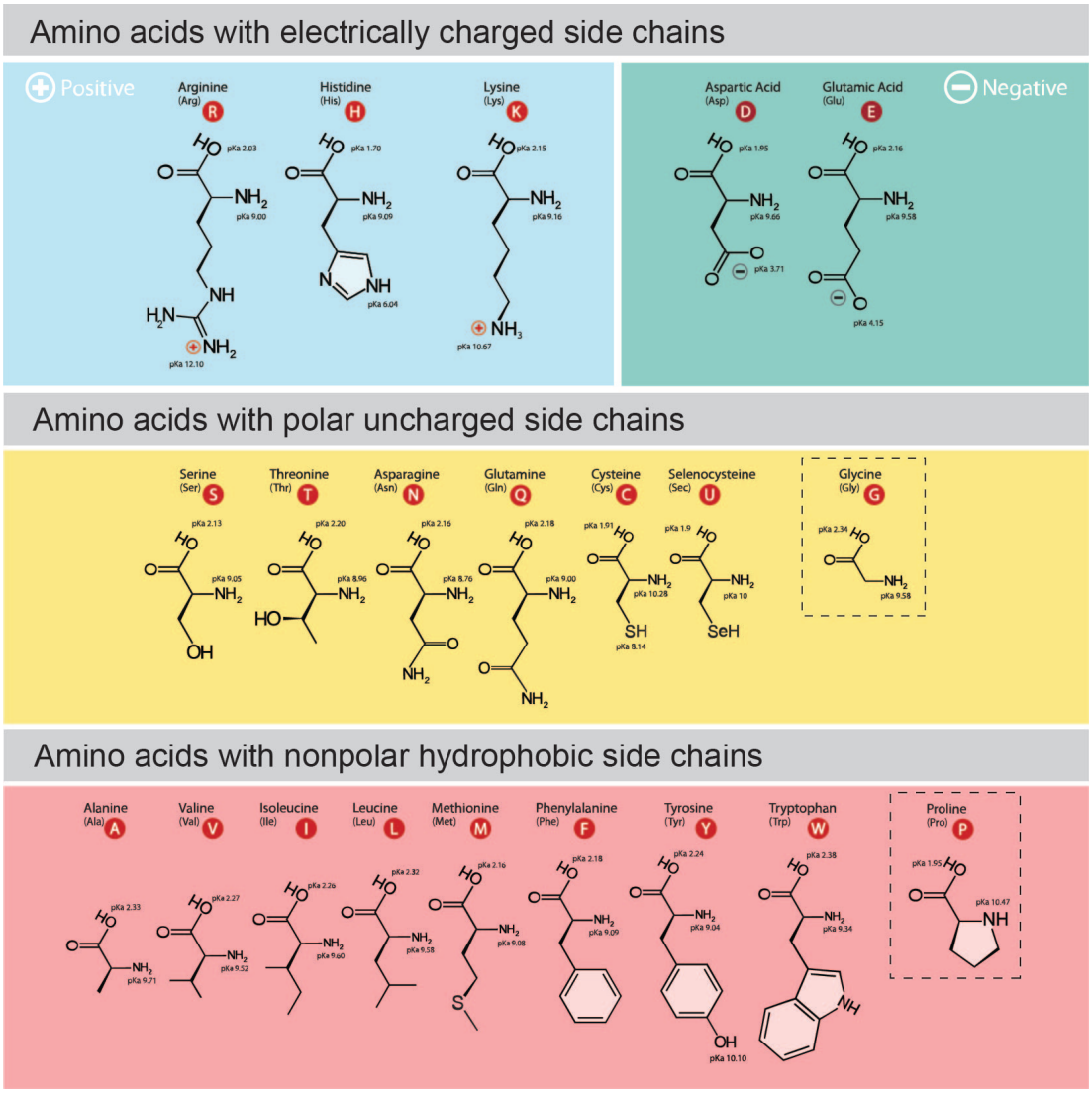


FIGURE 6.8: 21 amino acids, grouped by side chain characteristics. Residues in dashed boxes represent special cases that do not conform to the typical structure of the group. Letters circled in red indicate the amino acid shorthand labelling used throughout the analysis. (Modified from image by Dan Cojocari, CC License BY-SA 3.0.)

synthesised as an independent moiety, allows for affinity tests to be conducted prior to the expensive synthesisation process.

An overnight culture of ER2738 was diluted 1:100 in LB/Tet, sufficient for 20ml per clone to be amplified. The diluted culture was aliquoted 20ml per clone into labelled 125ml Erlenmeyer flasks, adding 10µl phage per flask before leaving to shake at

200rpm for 4.5 hours at 37°C. The amplification protocol described in 'Phage Amplification', Section 6.3.1, was then followed to produce amplified samples in 100µl TBS. Concentration was assessed using IPTG/Xgal plate titration as described in 'Eluate Titration', Section 6.3.1.

For the ELISA, primary target was added to sodium bicarbonate buffer to a total volume of 200µl, one well in a 96-well plate for each clone to be characterised. The target was left to bind overnight at 4°C. In the case of additional targets (such as purified p16 to antibody), the appropriate wells were aspirated and washed three times with TBS-T, the target added as described above and the plate incubated for 1.5 hours at room temperature. The plate was washed again in the same manner. All target wells and an additional row for block-control were filled with 200µl 5mg/ml BSA in $NaHCO_3$ solution. In addition, a separate 96-well plate was filled with BSA solution for phage dilution, with both plates incubated at 4°C for 1.5 hours. The sample and dilution plates were washed with TBS-T.

In the blocked dilution plate, 1:200 serial phage dilutions were carried out in 200µl TBS-T across four rows. 60µl of the final dilution (equivalent to ≈1010 virions) was added to the primary target row, secondary target row and BSA-only row for each phage to be tested, and the plate incubated with shaking for 1.5hr at room temperature. After washing, 100µl HRP-conjugated anti-M13 monoclonal antibody (GE Healthcare #27-9421-01, diluted 1:5000 in blocking buffer) was added immediately to each well to prevent drying and incubated with shaking at RT for 1hr.

After washing, sufficient ECL solution was prepared for 100µl per well. The solution was added quickly to the wells using a multichannel pipette and the plate immediately analysed at 425nm for 400ms on an Ascent Fluoroskan FL plate scanner (Thermo Scientific). The uncoated well luminescence values were averaged as an 'absolute blank' to indicate intrinsic measurement error.

Synthetic peptide sandwich ELISA

Peptides were synthesised from the picked DNA sequences in the format Biotin-SGSG-PEPTIDE-NH_2, where the peptides were 12 and 6 amino acids long for sequences against p16 protein (AN) and antibody (AB) respectively.

Streptavidin-coated 96 well plates were prepared by diluting a 2mg/ml Streptavidin (Sigma Aldrich SA101) stock 1:400 in dH_2O, adding 50µl per well and incubating at 37 overnight until dried. The plates were blocked with 100µl 3% BSA in PBS-T for 1hr at RT prior to addition of peptides. This ensured that binding was

streptavidin-biotin specific and therefore correctly oriented, rather than passive. Biotinylated peptides, synthesised by Mimotopes UK using Fmoc solid phase synthesis and supplied as a dry powder, were dissolved for 45min with shaking in 200µl 4:1 mix of DMSO/dH_2O to form a 25mg/ml stock solution, stored at -20°C when not in use. Stock solutions of both sample and control peptides (dissolved in 1ml dH_2O) were diluted in PBS-T before use. Following washes, 50µl peptide solution per well was incubated on the plates with shaking for 1hr at RT. The plate was washed again before addition of target to sample wells, again diluted with PBS-T. The sample wells were washed to remove unbound target before addition of 50µl relevant primary antibody to sample and control wells, diluted 1:1000 in PBS-T and incubated at RT for 1hr with shaking. Plates were washed before addition of HRP-conjugated rabbit anti-mouse secondary antibody (Polyclonal Rabbit Anti-Mouse HRP, DAKO) to all wells under the same conditions. Plates were washed before addition of 50µl ECL solution to each well. The plate was visualised immediately at 425nm for 400ms and the luminescence values analysed in triplicate against secondary-only control with blocked streptavidin-coated wells and true blank (empty wells) unless otherwise specified.

6.3.3 Phage samples for next-generation sequencing (NGS)

Phage plaques were picked in sets of 96 using sterile pipette tips and amplified in 1ml ER2738 culture each within a polypropylene 96-Well Deep Square Plates (2ml well volume, Thomas Scientific) at 200rpm for 4.5hr at 37°C. The plates were centrifuged at 3184g for 5min using microplate adapters in an Eppendorf 5810R centrifuge. The phage supernatant was transferred to new plates, sealed and stored at 4°C until further use. Each plate was pooled by taking 1µl from each well and amplifying the resulting volume using 1:100 ER2738 in 192ml LB/Tet in an Erlenmeyer flask at 200rpm for 4.5hr at 37°C. The resulting culture was aliquoted into 50ml Falcon tubes and precipitated as described in 'Phage amplification', Section 6.3.1. The amplified phage aliquots were then recombined for PCR, so that the sequences of individual phages could be analysed using next-generation sequencing.

A polymerase chain reaction (PCR) was carried out directly with the amplified pooled phage from each set alongside a phage-free control using UltraPure DNase-RNase-Free Distilled Water ('DNA-free dH_2O') in a 50µl reaction volume. The PCR reagents are detailed in Table 6.2; the reaction volume was prepared on ice to prevent premature denaturing of DNA. The primers, shown in Figure 6.9, were designed

to hybridise to the M13KE phage vector on either side of the randomised library insert, with additional adapters to be recognised by the Illumina MiSeq System as discussed below. The reverse primer was based on the 28gIII primer sequence, known to produce a readable sequence within a relatively short read from the insert of interest (Figure 6.7) as necessary for paired-end reads. The forward primer was designed on the basis of two limiting factors: the need for a suitable 'buffer zone' between the primer hybridisation site and the sequence of interest to ensure clean PCR amplification, and the presence of a forward terminator in the M13KE vector upstream of the insert. The forward and reverse primers were ordered from Sigma Aldrich, polyacrylamide gel electrophoresis (PAGE)- and high performance liquid chromatography (HPLC)-purified respectively. Primers were aliquoted once reconstituted and stored at -20°C.

FIGURE 6.9: The forward and reverse primers used to prefer the phage vector library for Illumina sequencing. The red sections of each sequence indicate the Illumina adaptors, which hybridise to the sequencing flow cell. The black regions indicate the phage-specific sequence. In each case the hybridisation region of the primer is shown in the context of the conserved phage sequence, underlined.

PCR was carried out using the thermal cycling conditions described below on a SureCycler 8800 (Agilent).

1. 95°C for 1min
2. 95°C for 15s
3. 55°C for 20s
4. 70°C for 1min
5. Steps 2-4 repeated 30 times
6. 70°C for 3.5min

Reagent	Supplier	[Stock]	Volume (Concentration)
Hercule II Reaction Buffer	Agilent (600675-52)	5x	10.00µl (1x)
Betain Solution for PCR	Sigma (B0300-1VL)	5M	5.00µl (0.5M)
D-(+)-Trehalose dihydrate	Sigma (T9531-5G)	100mM	6.6µl (13.2mM)
dNTP	Stratagene (200415-51)	100mM	0.25µl (500µM)
Primers (forward and reverse)	Sigma Aldrich	100µM	0.25µl (500nM)
Herculase II Fusion DNA Polymerase	Agilent (600675)	40 Reactions	1.00µl
Eluted pooled phage	In-house	-	1.00µl
DNA-free dH_2O	Invitrogen	-	25.65µl

TABLE 6.2: Constituents of buffers and other reagents used during PCR with pooled phage.

7. 4°C forever

The resulting amplified pool of phage vectors, prepared for next-generation sequencing with barcoded adaptors, were run on a gel to assess size and purity. Gels were produced by melting 1.5g Ultrapure Agarose in 100ml Tris-acetate-EDTA (TAE) buffer and adding 10µl SYBR Safe DNA Gel Stain (Invitrogen, S33102) before setting in a gel cast with sample comb. 50x TAE buffer stock was produced by dissolving 242g tris(hydroxymethyl) -aminomethane (Tris, VWR), 54.6ml acetic acid (Fluka) and 100ml 0.5M ethylenediamine tetraacetic acid (EDTA, pH8.0, VWR)) in 1l dH_2O. A 5µl aliquot of each sample was mixed with 1µl 6x Orange Gel Loading Dye (ThermoFisher R0631), loaded alongside 6µl DNA reference ladder and the gel run in TAE buffer at 100V for 45-60min depending on gel cast size. The gel was visualised using UV transillumination on a ChemiDoc MP Imaging System (BioRad) with Image Lab 6.0.1 software. DNA was extracted from the remaining sample volume using a QIAquick PCR Purification Kit (Qiagen 28104), following the kit protocol.

The final sample was sent for next-generation sequencing (Figure 6.10) on a MiSeq System at Edinburgh Genomics, with 11M, 250 base paired-end reads. A single-lane flow cell is loaded with the pooled samples, prepared with Illumina adaptors. The flow cell is populated with a lawn of two adaptor-complimentary oligonucleotides, one of which anchor the sequence library to the cell. A polymerase then creates a complement of the hydridised sequence. Bridge amplification is carried out, where sequences are tethered between the complementary flow cell adaptors before being separated. In this way, any sequences of the same length prepared with the correct adaptors can be amplified simultaneously. The reverse sequences are then removed.

Sequencing is carried out through the use of fluorescent nucleotides: as each labelled nucleotide is added to the growing sequences, the cell is illuminated with a light source and the nucleotide label recorded for each tethered chain. This process is then performed in reverse, generating two complementary sets of data.

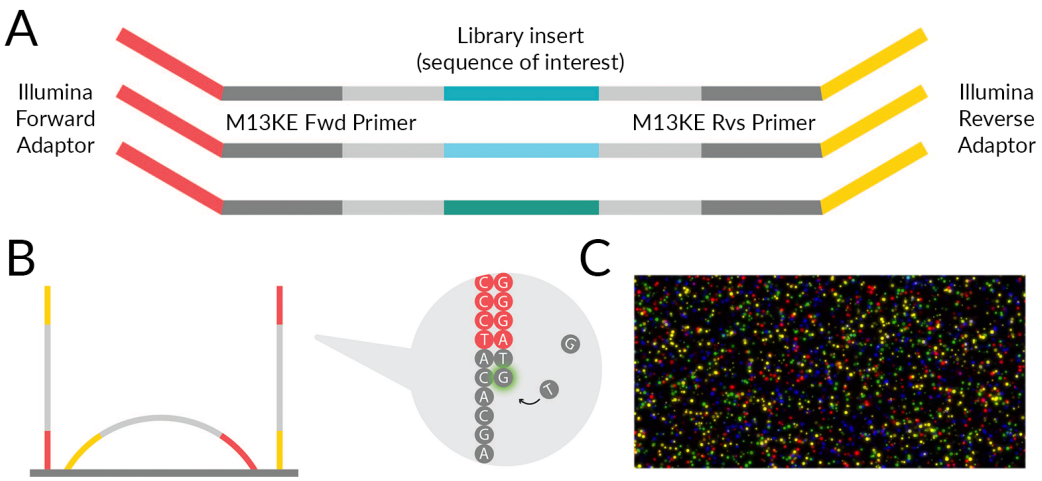


FIGURE 6.10: A schematic of the next generation sequencing process. A) The ready-to-sequence library following PCR amplification with Illumina-adapted primers, feating the Illumina forward (pink) and reverse (yellow) adaptors, the forward and reverse primers (dark grey), buffer sequence (light grey) and variable library insert (blues and greens). B) Bridge amplification of the sequences, which hybridise to the flow cell via the Illumina adaptors. Once released, the sequence 'read' takes place though addition of colour-coded bases. C) Image of the flow cell during one sequence cycle. The fluorescence emission is used to interpret the sequence of each tethered library.

6.3.4 NGS data analysis workflow

Owing to the small sample pool size relative to other NGS datasets, the samples were made up with 50% PhiX Control v3 Library by Edinburgh Genomics. The library consists of a well-characterised bacteriophage PhiX genome which provides more balanced fluorescent signals in the case of low sample sequence diversity. This sequencing data was removed by Edinburgh Genomics to leave 3.89M paired-end reads in compressed FASTQ Sanger format. FASTQ files contain additional quality information to standard FASTA files. Owing to the excessive number of reads for 192 samples, only the forward sequencing was analysed. Initial quality control analysis performed using FastQC, developed at the Babraham Institute [256]. This showed that the sequencing was high-quality and of the correct length (Figure 6.11).

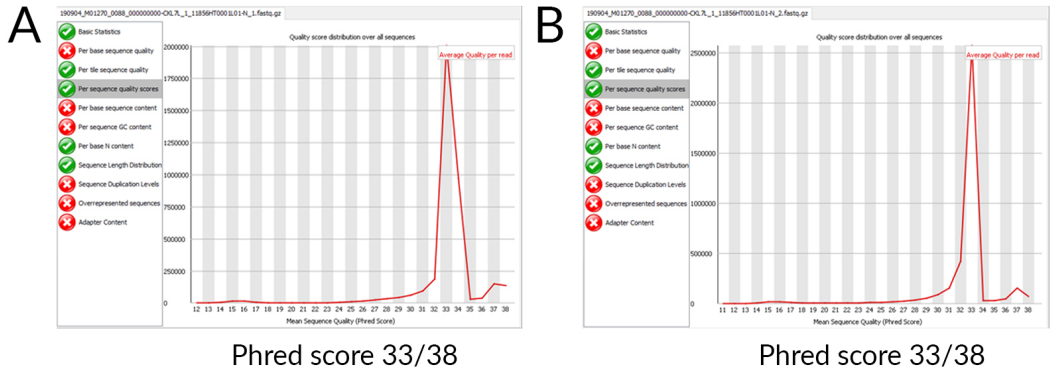


FIGURE 6.11: Quality control data for the forward (A) and reverse (B) reads from Edinburgh Genomics. An average Phred score per base of 33 is equivalent to 0.05% probability of an incorrect read. The modal read length was 250bp as expected, dropping off rapidly after 249 and 251bp (data not shown).

All pre-processing was performed using Galaxy, an online processing tool for large-scale sequencing projects. The forward-read file was decompressed and converted to FASTA format using Mothur v.1.39.5, developed by Patrick D. Schloss at the University of Michigan [257]. The forward adapter sequence was used as a Galaxy Clip guide, simultaneously trimming the M13KE phage sequences and removing sequences not containing the adapter. These are likely to be misreads.

The remaining 1,084,815 sequences were filtered to remove exact duplicates, reducing the file to approximately 1M reads. The first sequence was found to be modal and filtered out using the Galaxy Select tool to leave 210k sequences. These remaining sequences were subjected to 'Select' on the basis that they contain the characteristic glycine spacer associated with the phage insert, GGTGGAGGT. This left 9255 reads which were analysed as described in the results section.

6.4 Peptide discovery for p16 protein and antibody capture

6.4.1 Optimisation of the biopanning process

As described in Section 6.2, the lateral flow device requires peptide capture layers capable of binding to p16 (test line) and the conjugation antibody (control line). Moreover, the p16 capture layer should have affinity for the protein when already bound to the conjugation antibody. As such, two phage-displayed peptide panning experiments were carried out in parallel - one against p16 protein when bound to the conjugate antibody (labelled as 'AN') and one against the conjugate antibody in isolation

('AB').

Preliminary work with a three-round panning process as detailed in Section 6.3.1 yielded pooled phage populations demonstrating affinity for AN and AB targets with some phage library inserts in both pools showing affinity for BSA block, used at 30mg/ml (Figure 6.12). The amplified pools were titred on blue/white plates to yield separate phage plaques. AN and AB samples were found to contain 85×10^{11} and 129×10^{11} pfu/ml respectively. After some issues with efficient DNA precipitation from individual phage plaque-infected ER2738 cultures using the rapid sequencing protocol, 1ml rather than 0.5ml amplified phage samples were used.

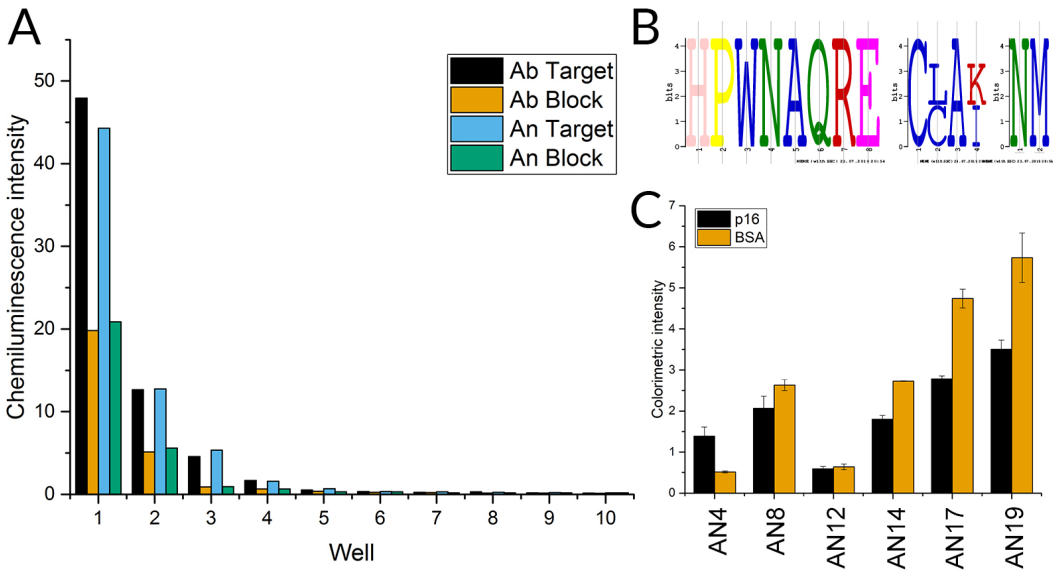


FIGURE 6.12: Characterisation of amplified Round 3 from the preliminary phage display optimisation. A) ELISA-based affinity testing of the phage pool, indicating the presence of binders for both targets as well as non-specific populations. B) MEMESuite analysis of the antigen sequences, indicating some consensus motifs. C) ELISA-based affinity testing of individual native phage.

18 AB samples and 20 AN samples were picked, amplified and prepared for sequencing. The sequencing results for AN and AB samples can be seen in Tables 6.3 and 6.4 respectively. AN2 is a shorter sequence, which can occur in the case of frameshift mutations in the phage vector. Sequences were analysed using SnapGene Viewer, with the randomised library insert identified using the characteristic bracketing sequences as described in 'Selection and characterisation', Section 6.3.1. Initially six sequences (AN4, 8, 12, 14, 17 and 19) were assessed for binding affinity on the basis of either consensus or similarity of amino acid sequence, analysed using MEMEsuite (Figure 6.12).

AN#	Amino acid sequence	AN#	Amino acid sequence
AN1	YVPWATLSQNSH	AN11	QDSTTLVPDEYI
AN2	DLLQKHSGWSA*	AN12	AETVESCLAKSH
AN3	ATTSIPPYWQAL	AN13	ALNWSPKLPVPP
AN4	NMDIEHSASFRW	AN14	HCCAITQNTDLA
AN5	NPKMSSNMAGPL	AN15	ATDWLSRYTYAA
AN6	TKNMLSLPVGPG	AN16	MPAVMSSAQVPR
AN7	LGNSPDFLVPTS	AN17	SHPWNAQRELSV
AN8	FSPLHTSTYRPS	AN18	QVIALARLPAPI
AN9	HYGMVMVRHIVA	AN19	SHPWNAQRELSV
AN10	HSSSASDRSRPL	AN20	HNARFPTSHPGG

TABLE 6.3: Sequences for the Round 3 antigen (AN) peptides individually sequenced as part of phage display optimisation. *AN2 is a shorter sequence by one residue. Amino acid shorthand labels are given in Figure 6.8.

Five of the six samples appear to have higher binding affinity for BSA than target.

Following this, the panning procedure was repeated with three key alterations to the initial method tested. First, the original randomised library was re-panned but with reduced BSA concentration (5 instead of 30mg/ml) at the blocking and washing stages to reduce the likelihood of selecting BSA-binding sequences. Second, four rounds of amplified panning were completed to further enrich the pool for high-affinity peptides. Finally, sequencing was performed in parallel to ELISA-based affinity testing. This is due to the fact that the preferential amplification of certain sequences over others, leading to consensus sequences within a cultured pool, is due to the ability of individual phage sequences to replicate and is not related to their binding affinity for a given target.

6.4.2 Discovery of peptide sequences

Prior to repeating the panning process, the library was reassessed for activity. IPTG/XGal plating of the unpanned library provided successful ER2738 infection and yielded a titration of approximately 1×10^{11} pfu per 10µl when averaged across ten-fold serial dilutions from 10^9 to 10^{12}. Following the first round of panning, the titrations for the unamplified AN and AB pools were 11×10^4 and 21×10^4 pfu per 10µl.

The panning process was then carried out as described and modified previously. Following amplification of the fourth round, the phage pools were titred on IPTG/Xgal plates and 24 plaques picked from each pool for an initial amplification, carried out in

AB#	Amino acid sequence	AB#	Amino acid sequence
AB1	SAHQLLLNKMPN	AB10	GSTQAWMSPPLA
AB2	TMGFTAPRFPHY	AB11	ATWSHHLSSAGL
AB3	TMGFTAPRFPHY	AB12	NTYVPNLLSSFR
AB4	DSTLHKNRLQTR	AB13	DQMPPLNRSPSL
AB5	GSDLYASKIWTA	AB14	AEQGTYRVWLLP
AB6	RPSYTDHDHRSN	AB15	SWWPFPPQPDPA
AB7	YASSTKADGFLS	AB16	YQLRPNAESLRF
AB8	TQCCSSGATYPR	AB17	THSKSTSLWDSN
AB9	FVSTTMKLSADS	AB18	QSFASLTNPRVL

TABLE 6.4: Sequences for the Round 3 antibody (AB) peptides individually sequenced as part of phage display optimisation. Amino acid shorthand labels are given in Figure 6.8.

3ml LB as described in 'Selection and Panning', Section 6.3.1. A further amplification was performed in 20ml cultures before ELISAs were carried out.

Owing to the large number of samples, titration of the amplified phage stocks was carried out for a representative selection of samples and the individual clones diluted on this basis. AN1, 7, 13 and 19 yielded titrations of 1.0×10^{20}, 4.6×10^{19}, 5.7×10^{19} and 1.8×10^{19} pfu per 10µl respectively. As such, only intra-sample comparison can be made from native phage ELISAs, as luminometric intensity is partly dependent on these small variations in sample concentration. As described in 'Native phage peptide affinity ELISA', Section 6.3.2, phage samples were diluted 1:200 in TBS-T using the blocked dilution plate, to yield sample titres for incubation on the order of 1.25×10^{10} virions per well.

Antigen samples were affinity tested in two formats. Both used p16 bound to BD Pharmingen antibody as in the original panning procedure, and BSA only as a blank control. Final concentrations per well of 0.5ng/µl p16 mAb and 100ng/µl purified p16. The first assay used p16 only as a positive control, and the second assay antibody (mAb) only, as a negative control. This allowed for consideration of specificity and affinity, particularly to account for the selection of phages specific to antibody rather than target. Samples had to be assessed in singlicate owing to the limited availability of purified target and the large number of samples.

For the first assay (Figure 6.13A), it was posited that samples which bound to p16+mAb and BSA at the same or higher affinity than p16 only were unlikely to be specific for the target. The second assay (Figure 6.13B) was carried out to confirm binding affinity for p16 without the issue of randomly oriented target and potential denaturation upon passive binding to the plate. Here, samples producing a higher

luminescence intensity for mAb or BSA than mAb+p16 were judged as non-specific. However, it is worth noting that there are caveats to using a native phage-based ELISA approach.In addition to non-specific binding arising from the peptides themselves, the phages can also adhere to the antibody, protein or complex thereof, as well as the BSA blocking layer.

The results for the AN sample ELISAs can be seen in Figure 6.13. Figure 6.13A suggests, as expected, that the random orientation and potential denaturation of p16 upon plate binding reduces, in some cases heavily, the ability of the phage peptides to bind to their target in a different conformation to the one in which they were selected. Some samples show a higher BSA than target affinity, indicating non-specificity. The potential candidates from this assay - those demonstrating higher p16 than BSA luminometric signal - are samples AN1, 3, 9, 10, 15, 17, 19, 21 and 22.

For the negative control test in Figure 6.13B, the samples producing a higher signal for mAb+p16 than mAb only are AN14, 16, 20, 21, 22, 23 and 24. Sample AN20 appears to be binding to BSA. AN1, 3, 9, 10 and 17 appear to be specific for mAb, leaving only AN22 as a cross-referenced potential candidate in this context.

The affinity ELISA for AB samples can be seen in Figure 6.14. Here, mAb+p16 was added as a negative control to compete out peptides with specific affinity for the binding site of the antibody.Firstly, it is clear that there are general issues with binding strength of Samples 13-24 which may be the result of the original panning process but is more likely to result from issues with solution volume error arising from the multi-channel pipette used for the assay, probably during the sample dilution phase. Samples AN1, 2, 7 and 12 indicate high levels of p16+mAb binding, rendering them unsuitable for discovery of the p16 epitope or for use as a paper-based assay control peptide.

The samples were sequenced according to the rubric described in Section 6.4.1 - the results for AN and AB samples are detailed in Figure 6.15. Samples 8, 14 and 17 had no insert. Sample AB4 did not yield a readable sequence after multiple sequencing rounds, and AN2 lacked a glycine spacer and contained non-library codons, indicating wild-type phage. Many of the AN samples are consensus or near-consensus sequences, as expected after four rounds of panning. However this data also demonstrates that some of the conserved sequences produce different affinity results with native phage ELISA, not simply by intensity but also by ratio. This may well be due to the caveats described earlier. It is also possible that a combination of lack of replicates and low target density, both arising from lack of viable target availability, mean that binding affinities are being assessed very close to the limit of detection and are therefore subject to low signal-to-noise ratio. In the future, the native phage ELISA could

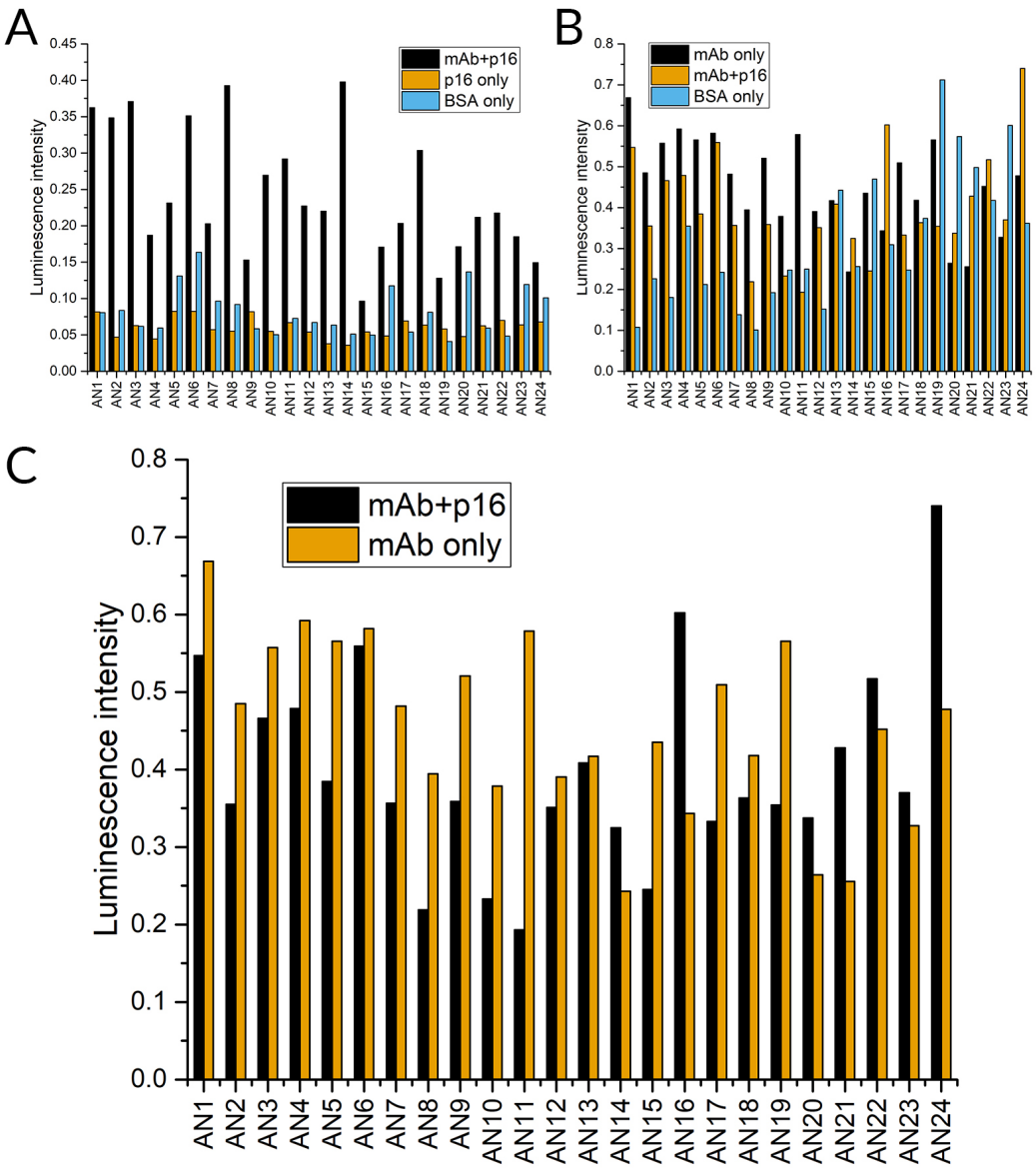


FIGURE 6.13: ELISA affinity results for amplified Round 4 AN phages. A) Phages tested against original panning target (mAb+p16), target only (p16) and BSA control. B) Phages tested against original panning target (mAb+p16), negative (mAb only) and BSA controls. C) Repeat of graph B without BSA results for ease of viewing.

be completely replaced with an assessment of peptides synthesised on the basis of sequencing data, or in the instance of large datasets, used to make a smaller selection of phage when conducted in triplicate.

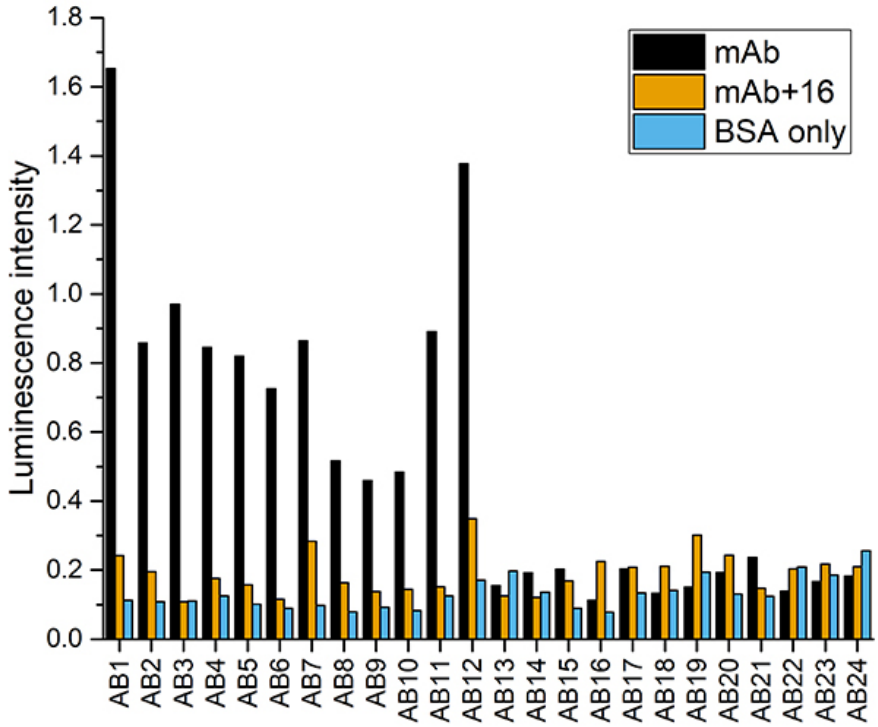


FIGURE 6.14: ELISA affinity results for amplified Round 4 AB phages, tested against original panning target (mAb), competition control (mAb+16) and BSA control. Only one sequence (AB2) was found to have similarity to p16 protein sequence.

6.4.3 Affinity of synthetic peptides by ELISA

Peptides were synthesised as described in Section 6.3.2. For AN samples, each unique sequence (Figure 6.15) was produced. For AB samples, the amino acid sequences were compared to p16 protein sequence in order to find a potential motif. Only one sequence, AB2 (W—FYPGRIPG), was found to be a potential match for the p16 sequence (WGRLPV). To test for potential affinity for the antibody being used, an alanine scan was carried out as seen in Table 6.6. The synthesised peptides were tested for affinity in a sandwich ELISA as described in Section 6.3.2. The antigen samples required both target and antibody to be introduced; for antibody samples, only the latter was required. For the ELISA results shown in Figure 6.16, the 'primary target' (p16 protein or antibody) was incubated at 4°C overnight to ensure binding.

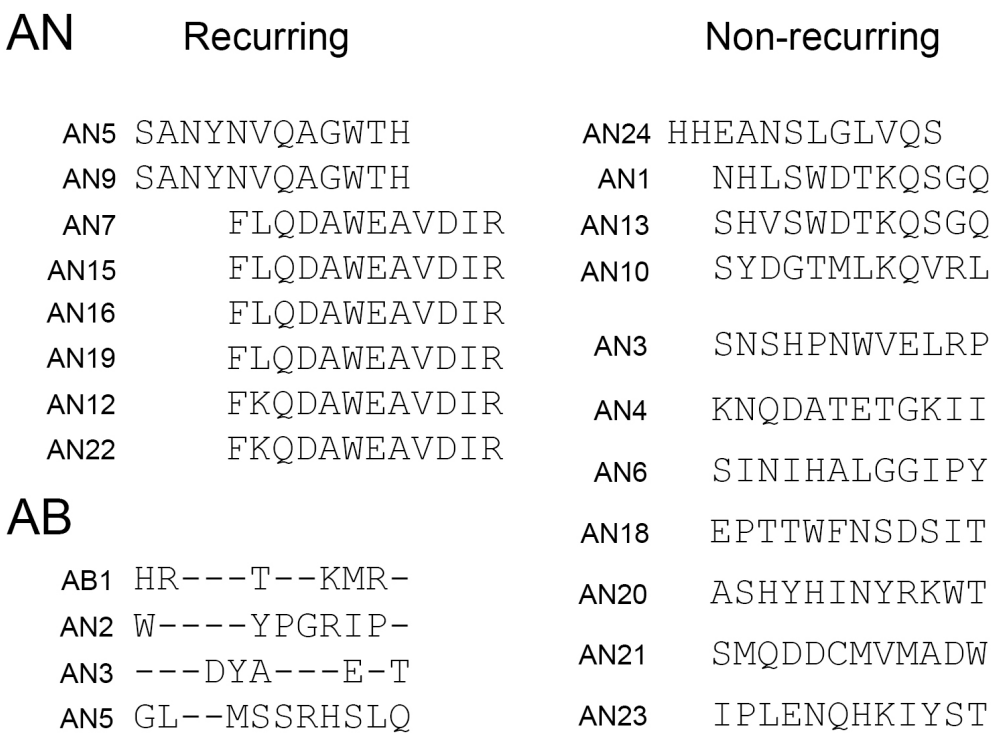


FIGURE 6.15: Peptide sequence summary for Round 4, with antigen samples organised according to recurring or non-recurring sequences. Similar sequences are provisionally aligned. Antibody samples were sequenced using a different supplier, and returned poor-quality sequence information.

The control peptides were provided by Mimotopes as a guide for assay development rather than direct lumimetric comparison and so are shown separately. The controls used both different working dilutions (1:10 as opposed to 1:1000) and a different antibody. The 'positive' and 'negative' control sequences and expected signal intensity from a control antibody titration can be found in Figure 6.17. The small difference in amino acid variation accounts for the relatively high absorbance signal produced by the nominally negative control peptide, but also demonstrates the small sequence size necessary to create a strong binding affinity.

Table 6.5 lists the antigen peptide samples and their corresponding phage sample number in order of luminometric intensity, with samples significantly higher ($p < 0.01$) than secondary-only control displayed above the dividing line. The results show some consistency with the second native phage ELISA (Figure 6.13C). Intra-sample analysis shows that all of the phage samples demonstrating a higher signal for p16+mAb

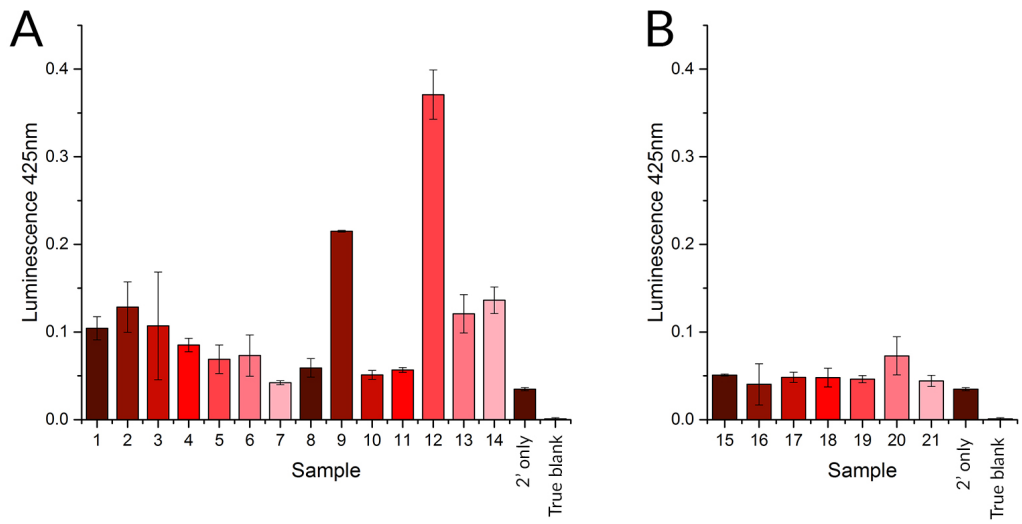


FIGURE 6.16: Sandwich ELISA affinity of synthetic peptides, both tested in triplicate at 1:1000 stock dilution. A) Antigen (AN) peptides, tested against purified p16. B) Antibody (AB) peptides, tested against monoclonal antibody. Secondary only control contained target but no primary antibody.

over mAb only (An14, 16, 20, 22, 23 and 24) feature in the statistically significant synthesised list in Table 6.5. The second highest affinity synthesised sequence (Sample 9, equivalent to phage sample An6) does not feature at all in the phage ELISA 'hits', however. Upon sequencing, An7, 15 and 19 were found to be identical to An16 and yet produce different binding characteristics in native phage ELISA, again indicating the error-prone nature of the first affinity assay tested.

Analysis of the amino acid groupings and potential motifs are considered in Section 6.4.4.

The antibody peptide samples are shown in Table 6.6. These are not listed in order of intensity, as none of them produced a signal statistically different to that of the secondary-only control at the same level as the antigen samples ($p < 0.01$). This would indicate that the six-residue sequence is not sufficient for binding on its own, as the remaining panned sequences did not share any common regions with the p16 protein sequence.

The four highest affinity samples were then tested against p16 titrated in a serial dilution (Figure 6.18), with Sample 12 demonstrating increased binding sensitivity across all concentrations. The peptide sequence is ASHYHINYRKWT. Controls for the samples included p16 only and mAb only samples to test for specificity. Both the sample and control peptides were used at 1:100 dilution. The results are consistent

An#	Mimotope Sample	Sequence	Molecular weight (kDa)
An20	12	ASHYHINYRKWT	2.089
An6	9	SINIHALGGIPY	1.768
An23	14	IPLENQHKIYST	1.956
An7, 15, 16, 19	2	FLQDAWEAVDIR	1.976
An21	13	SMQDDCMVMADW	1.945
An12, 22	3	FKQDAWEAVDIR	1.991
An5, 9	1	SANYNVQAGWTH	1.860
An24	4	HHEANSLGLVQS	1.804
An13	6	SHVSWDTKQSGQ	1.873
An1	5	NHLSWDTKQSGQ	1.914
An3	8	SNSHPNWVELRP	1.949
An18	11	EPTTWFNSDSIT	1.911
An4	10	KNQDATETGKII	1.831
An10	7	SYDGTMLKQVRL	1.924

TABLE 6.5: The synthetic peptides against p16, produced from phage sequences discovered through the biopanning process. Sequences displayed above the dividing line produced significantly stronger luminometric signal than secondary-only control.

Mimotope Sample	Sequence	Molecular Weight (kDa)
15	**WGRLPV**	1.240
16	WGRLP**A**	1.212
17	WGRL**A**V	1.214
18	WGR**A**PV	1.198
19	WG**A**LPV	1.155
20	W**A**RLPV	1.254
21	**A**GRLPV	1.125

TABLE 6.6: The synthetic peptides against p16 antibody, produced from phage sequences discovered through the biopanning process. Each sample represents an alanine scan of the proposed p16 binding site (Sample 15).

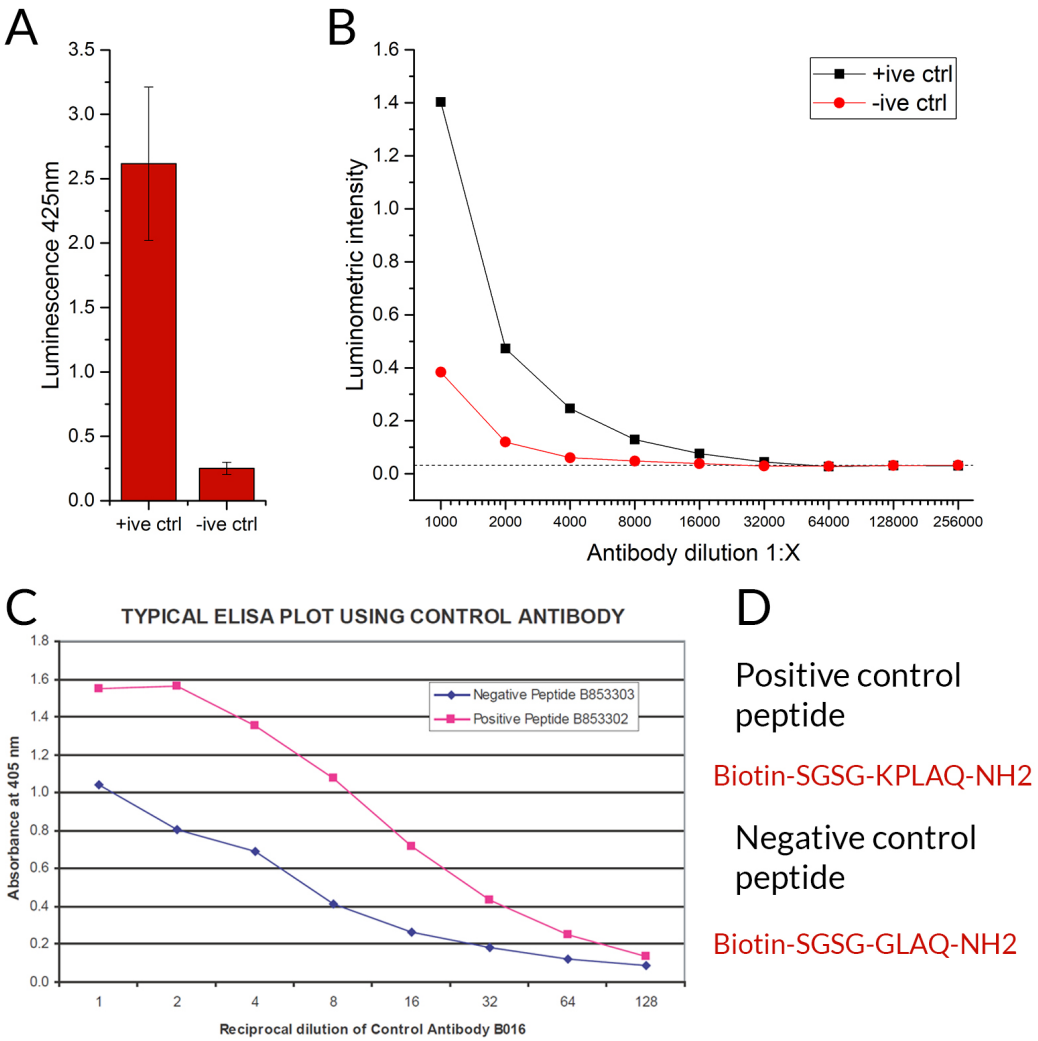


FIGURE 6.17: Characterisation of control peptides supplied by Mimotopes. A) Sandwich ELISA affinity, tested in triplicate at 1:10 stock dilution alongside samples in Figure 6.16. B) Repeat sandwich ELISA at 1:100 stock dilution with antibody titration. C) Reference ELISA data supplied by Mimotopes, using 1:10 stock dilution. D) Sequence of control peptides.

with those shown in Figure 6.16 in terms of binding affinity. The increase in signal with p16 concentration increase, while the antibody content is kept constant, indicates specificity for target.

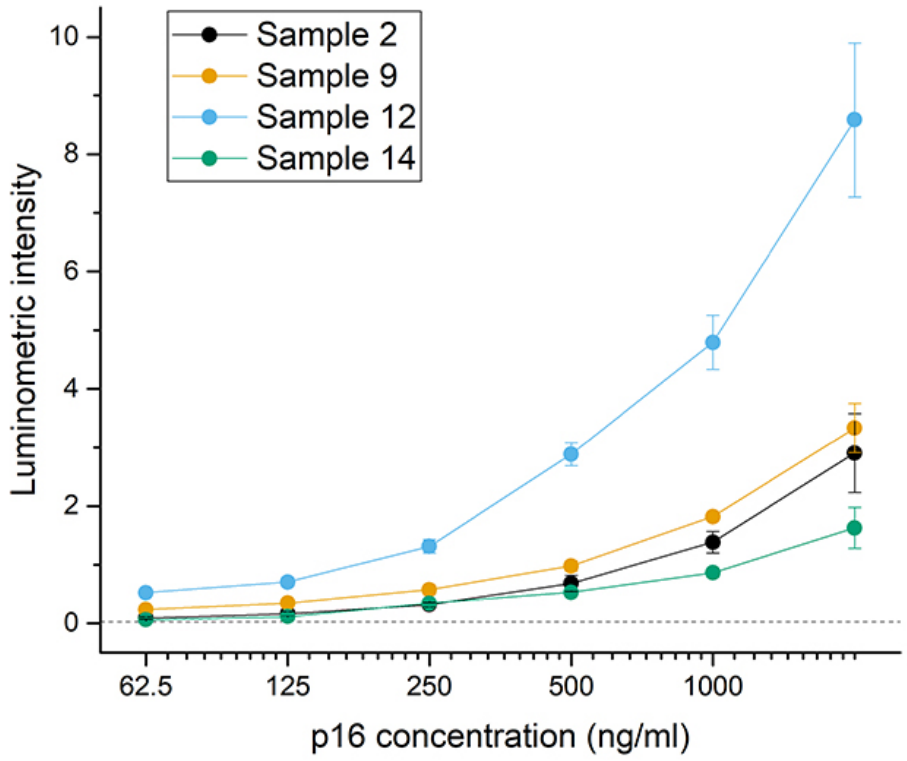


FIGURE 6.18: Serial dilution curve of purified p16 against a 1:100 stock dilution of the four highest affinity peptides from Table 6.5, tested in triplicate.

6.4.4 Amino acid analysis of peptide sequences

Due to the randomised nature of the phage insert library, it is important to analyse the amino acid residues for consensus sequences and potential epitopes for the target of interest. Initially, the unique sequences shown in Table 6.5 were uploaded to GLAM2: Gapped Local Alignment of Motifs, a sub-programme of MEMESuite capable of aligning sequences with potential gaps in the motif. In hindsight the gap functionality was not required but this initial step still allowed for rapid alignment and the identification of modal sequences at each position.

The full sequences were then considered as summarised in Figure 6.19 and potential alternatives added manually (shown in italics). This was justified on the basis that repeated or shared motifs within the same sequence may well produce a positive binding assay signal from two different binding configurations. On an unweighted

basis (without considering peptide affinity), the modal residues in the 'core' region (positions 9-13) are N(7,0), Q(5,5), D(5,1), A(4,0) and W(8,0). The values in brackets correspond to the frequency of the given amino acid in the GLAM2 and manual considerations respectively. In addition, the most frequently occurring residues (N at position 07 and W at position 11) feature alongside similar amino acids from other sequences, namely S and L/I respectively.

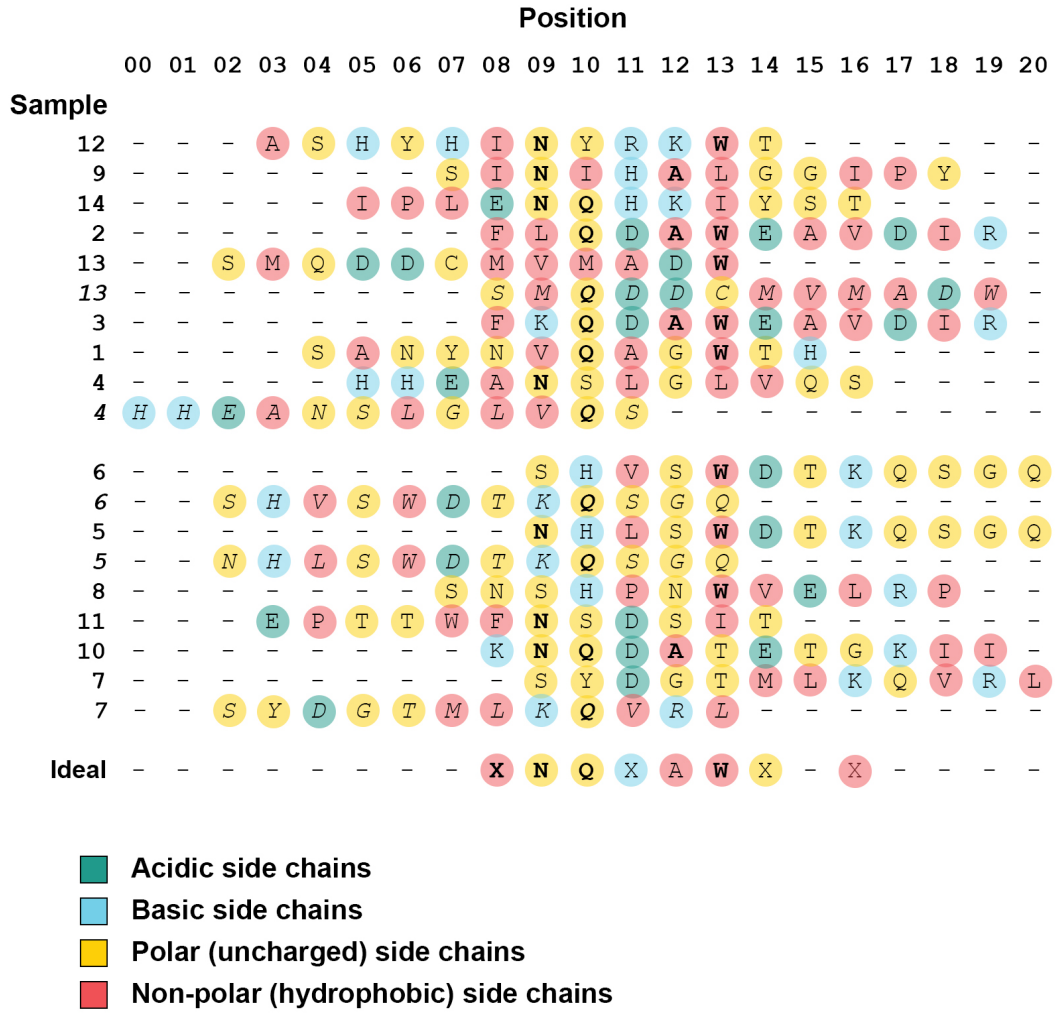


FIGURE 6.19: AN peptide sequence alignment and residue analysis, summarised as an idealised motif. Sequences are ordered by affinity and residues highlighted according to their grouping.

The motifs were then ordered by binding affinity to purified p16 as previously listed in Table 6.5. This was to increase likelihood of highlighting the true motif, rather

than an ineffective but consensus sequence from the inital randomised library. This highlighted the potential importance of different amino acids in creating a specific binding sequence. As before, the N, Q and W residues appear to be important, but only in the context of neighbouring basic and non-polar regions. Based on frequency and order, there also seem to be contributions from hydrophobic residues at positions 08 and 16, and to a lesser extent polar residues at position 14.

Analysis of positions 11 and 12 becomes more complex when the sequences are ordered by affinity. Basic residues - those positively charged in a pH neutral environment - are not especially common overall and yet appear in the top three hits in this region. Owing to the heterogeneous population at position 12, alanine (A) was chosen as being the least disruptive residue due to its shared structure with the majority of other amino acids.

The tentatively assigned 'Ideal' sequence, also shown in Figure 6.19, offers opportunities for further optimisation of sequence affinity which was beyond the scope of this thesis.

6.5 Peptide activity with cell models

6.5.1 Quantification of p16 in cell lines

Following the emergence of a high-affinity sequence for p16 when bound to antibody (ASHYHINYRKWT), consideration of the peptide's efficacy in a cell model context was considered. First, the ability of the purified target to be detected in the context of cellular lysate was tested by Western Blot. This was to explore the extent to which the epitope may be cleaved or denatured in the presence of lysis buffer or other active cellular components. A purified protein ladder was also produced as a quantifiable reference for the level of p16 present in the cell models used. The results can be seen in Figure 6.20.

Figure 6.20A and B demonstrate a serially diluted reference ladder of purified p16, introduced to the wells alone and spiked into HaCaT lysate respectively. The remaining 20µl sample volume was made up with dH_2O before being mixed with sample buffer and denatured by heating, as described in Section 2.3.1 ('Bradford Assay') of Chapter 2. The reference ladders themselves demonstrate that purified p16 is preserved in the presence of HaCaT lysate (p16 undetectable) and through the Western Blotting process, although the protein is dimerising to a greater extent in this context. As little as 0.79ng absolute p16 protein content could be detected using a 20s exposure, equivalent to approximately 0.0001% of total protein content or 40ng/ml.

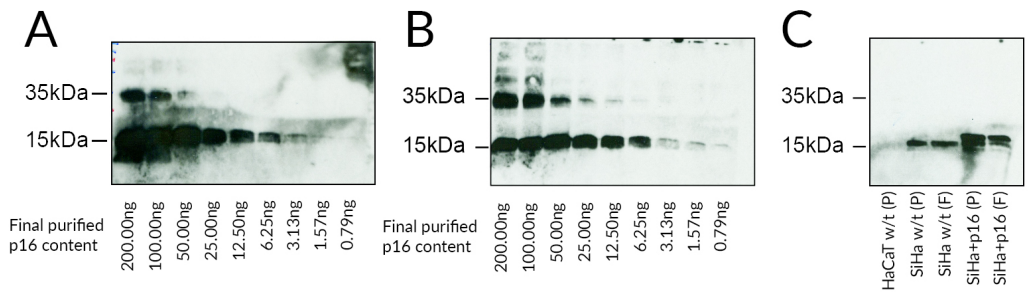


FIGURE 6.20: Western Blots demonstrating p16 expression. A) Purified p16 reference ladder in dH_2O. B) Purified p16 reference ladder spiked into HaCaT lysate. C) Cell lines used for p16 binding affinity of peptides, including HaCaT w/t (p16 undetectable due to hypermethylation of promoter [258, 43, 259]), SiHa w/t (high endogenous p16) and SiHa+p16 (high endogenous and transfected his-tagged p16). SiHa samples were tested in both pelleted snap-frozen (P) and fixed (F) formats.

Sample	Format	Absolute content	Percentage	Concentration
HaCaT w/t	Pellet	0ng	0%	0ng/ml
SiHa w/t	Pellet	3-6ng	0.05-0.1%	150-300ng/ml
SiHa w/t	Fixed	3-6ng	0.05-0.1%	150-300ng/ml
SiHa+p16	Pellet	6-25ng	0.1-0.4%	300-1250ng/ml
SiHa+p16	Fixed	6-12ng	0.1-0.2%	300-600ng/ml

TABLE 6.7: Expression of p16 in cell line models as interpreted from the p16 in lysate reference ladder shown in Figure 6.20B. Expression shown in terms of absolute content, percentage of Western Blot sample volume, and sample concentration.

The cell lines used were compared to the reference ladder in HaCaT lysate to assess approximate levels of p16 expression. The results are summarised in Table 6.7. When compared with Figure 6.18, it can be seen that the protein concentrations are well within the operating range of the four highest affinity peptides and particularly Sample 12.

6.5.2 Peptide binding efficacy with cell lysate

The best performing peptide (Sample 12, ASHYHINYRKWT) was tested in the presence of lysate, reserved from the samples demonstrated in Figure 6.20. This was both to investigate the ability to differentiate between levels of p16 expression in cell models rather than through an idealised pure protein ladder, but also to confirm that the peptide's biological functionality is not disrupted in the presence of crude cell lysate.

The ELISA was carried out according to the protocol described in 'Synthetic peptide sandiwch ELISA', Section 6.3.2, with three instead of five washes following lysate addition to mitigate for lower levels of target protein which might otherwise be removed altogether. Prior to addition, the samples were heated to 70°C for 4min as in Western Blot sample preparation so that the results were comparible. The samples were tested with the same absolute protein content as their Western Blot counterparts in Figure 6.20, made up to 50µl with PBS-T to ensure well coverage. The samples were tested with two technical replicates, with the luminometric response shown in Figure 6.21.

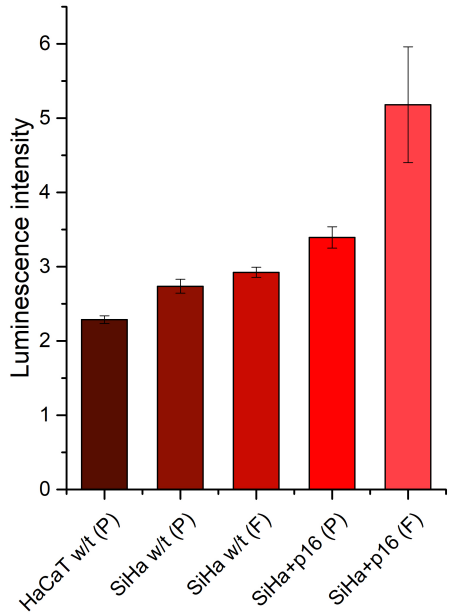


FIGURE 6.21: ELISA-based capture of endogenous and transfected p16 by synthetic peptide in a sandwich format. Luminometric results were visualised using BD Pharmingen primary monoclonal antibody to p16 and polyclonal HRP-conjugated secondary. SiHa models were analysed in both pelleted snap-frozen (P) and PreservCyt-fixed (F) formats.

In terms of the cell lines used, the results are consistent with the Western Blot in Section 6.5.1 - the assay demonstrates the lowest response for HaCaT wildtype, a middling response for SiHa wildtype and highest response for SiHa+p16. However, the response to 'native' snap-frozen, pelleted versus PreservCyt-fixed cells differs with this format. A higher signal is seen in cells which were stored in fixative before lysing. It is possible that methanol fixation better preserves the p16 binding epitope for the peptide, which is a positive outcome for application to clinical smear samples.

There is a significant difference ($p \leq 0.05$) between each of the sample types but not between the same sample types processed in pelleted or fixed form. This indicates that the chosen peptide and antibody could be used in conjunction to assess p16 expression in PreservCyt-fixed cervical cells. The control well, consisting of primary and secondary antibody incubated with streptavidin-coated, BSA-blocked plate, demonstrates that the assay is prone to variability from the visualisation method alone regardless of peptide affinity or lysate content. This may be due to non-specific interactions with BSA as seen in Chapter 4, or because of the less stringent washing conditions used. To improve background variation, the initial washing conditions (5x) should be adopted for all steps.

6.6 Next generation sequencing of phage populations

6.6.1 Large-scale sequencing of phage populations

Following the residue analysis carried out in Section 6.4.4, a larger population of phage plaques were selected by blue/white plate titration from the amplified fourth round pool. To increase the likelihood of identifying high-affinity but low-frequency motifs, 192 separated phage populations were prepared for sequencing in two pools of 96 (A and B), as described in Section 6.3.3.

The use of next generation sequencing allowed for the phage populations to be amplified and sequenced as a pool, thereby negating the requirement for batch processing of small culture volumes. The purified, pooled phage vectors were amplified directly with Illumina adaptors alongside dH_2O control as described in Section 6.3.3. 5µl of each sample was run in a DNA gel alongside 10bp O'Range Ruler and 1kb DNA ladder (NEB) to assess purity and size (Figure 6.22) - the gel was visualised using UV trans-illumination with 1.227s acquisition, optimised for faint bands. The ready-to-sequence pools were approximately 250bp in length as expected. The full product, as illustrated previously in Figure 6.10, is 251bp including Illumina adaptors. Any wild type phage contamination in the PCR would have resulted in an insert-free population at 203bp and therefore an additional band, which was not seen.

The remaining 45µl of each pool was purified and sent for analysis on an Illumina MiSeq platform (Edinburgh Genomics), using 50% PhiX volume due to the similarity of the sequences. The individual amplified pools A and B yielded concentrations of 94.1 and 122.5ng/µl respectively. The pools were combined to yield a ready-to-sequence pool of 108.3ng/µl, equivalent to 637.7nM.

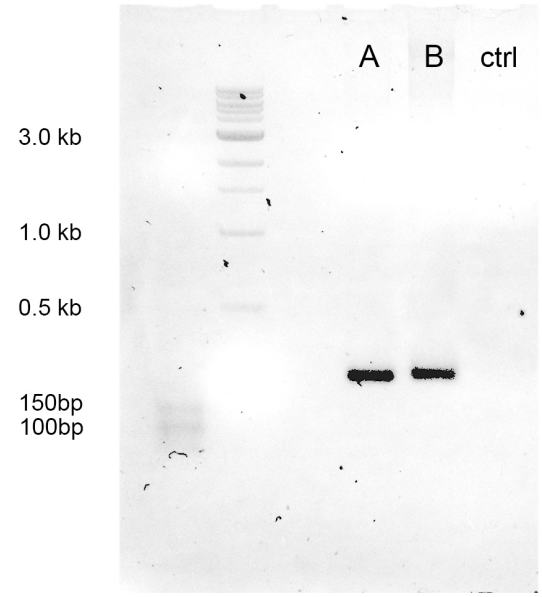


FIGURE 6.22: Band visualisation of the two phage pools following PCR amplification with Illumina adaptors, with a dH_2O-only negative PCR control to confirm primer orthogonality.

6.6.2 Analysis of large-scale peptide sequencing

Following removal of the PhiX sample data by Edinburgh Genomics, the NGS dataset was analysed primarily using Galaxy as described in Section 6.3.4. Following initial processing, the remaining 9255 reads, consisting of distinct sequences and their replicates along with nonsense reads, were analysed in Excel. The 'Find & Replace' function was used to highlight banks of identical sequences, thus reducing the need to analyse individual reads. The most common 192 sequences are summarised by order of frequency in Table 6.8 - only the peptide insert is shown. It was assumed that the remaining unique sequences (200+), all with a frequency of 1, were misreads of other sequences as only 192 unique samples were submitted as a pool.

TABLE 6.8: Phage peptide insert sequences and their associated frequency.

Sequence	Freq.	Sequence	Freq.	Sequence	Freq.
QAHDVMRGRMPG	860k	QAHDVMRGRMPG	55	QAHDVMRAYAG	4
QAHDVMRGRMPG	745	QALDVMRGRMPG	55	PAHDVMRGRMPG	4
QAHDVMRGRMSG	423	QAHDVMRGLMPG	53	QAHEVMRGRMPG	4
QAHDVMRGRMPG	398	QAHDAMRGRMPG	52	QAHAVMRGRLPG	4

QAHDVMWGRMPG	240	QAHDVMRGRMPV	52	QAPAVLRGRLPG	4
QAHDVMRGRMPG	231	QAHDVMRGRMPG	48	QAHDVMRGRRPG	4
QAHDGMRGRMPG	194	QAHYVMRGRMPG	48	QAHDVMRGGCR	3
PAHDVMRGRMPG	192	QAHDVMRGGMPG	46	QAQDVMRGRMPG	3
QAHDVMRGVCR	191	QAHDVMRGRMPD	44	QAHDVMRGRMPA	3
PAHDVMRGRMPG	186	QAHDVMRGRMPS	44	QAHDVMRGRMR	3
QAHDVMRGHMPG	130	QAHDVTRGRMPG	39	QGMMLCGGVCR	3
QAHDVMRGRMLG	128	QAHVVMRGRMPG	39	HAHYFIRGRIPG	3
QARDVMRGRMPG	119	QAHDVMRGRIPG	38	QAPDVMRGRMPG	3
QAHEVMRGRMPG	118	QAHDVCGGVCR	37	QAHDVRRGRRPG	3
QAPDVMRGRMPG	112	LAHDVMRGRMPG	36	QAHDVMRGGMRG	3
QAHDVMLGRMPG	110	EAHDVMRGRMPG	35	QSHDVMRGRMPG	3
QVHDVMRGRMPG	108	QAHDVMRGRKPG	31	PAHAVMRGRMPG	3
QAHDVMRGRMPG	104	QADDVMRGRMPG	31	QEHDVMRGRMTG	3
QAHDVMRGRIPG	103	QAHDFMRGRMPG	31	QAHYVIRGRMPG	2
RAHDVMRGRMPG	99	QAHDVMRGRMRG	26	QAHDGMRGRRPG	2
QAHDVMRGRMPG	99	QGHDVMRGRMPG	26	QAHDVMLGRMPG	2
QAHDVMQGRMPG	98	QAQDVMRGRMPG	25	HAHYVIRGRIPG	2
QAHDVMRGRMQG	98	QAHDVMRGRMPG	24	QAHEVMGGRMPG	2
QAHDVMRWRMPG	96	QAHDVMGGRMPG	23	QAHDVMRGGFP	2
QAHGVMRGRMPG	96	QAHDDMRGRMPG	22	QAHDVKRGSMRG	2
QAHDVMRGRMPG	95	QAHDIMRGRMPG	20	QAQEVMRGRMPG	2
QAHDVMRGRMPG	92	QAHDVLRGRMPG	20	QARDVMRGRMPG	2
KAHDVMRGRMPG	90	QAHDVMRGRMPG	19	QARDVMRGRMPG	2
QAHDVRRGRMPG	86	QAHDVMRGRMAG	19	QAHDFMRWRMPG	2
QAHAVMRGRMPG	86	QAHDVMRGRLPG	18	QAHDVMRGRMPG	2
*AHDVMRGRMPG	83	QANDVMRGRMPG	16	QAHDVMRGRMKG	2
QAHDVMRGRMPC	83	QAHDVMRGRMPR	15	QAHYVLRGRMPG	2
QAHDVMRGRMTG	82	QAHDGMRGRMPG	13	PAQDVMRGRMPG	2
QAHDVMRGSMPG	80	QAHDVMGGVCR	12	QAHDGMRGRMRG	2
QAHDVLRGRMPG	79	QAHDVMRGAYAG	12	QAHDGMRGRMPG	2
QAHDVIRGRMPG	79	QAHDVMRGRIPG	12	QAQDVMGGRMPG	2
QAHDVMRGRLPG	77	QAHDVMRGRMPG	12	QAHDVMRGLMPC	2
QAHDVMRGRRPG	76	QAHDVMRARMPG	12	PAHDVLRGRMPG	2
QAHDVMRGRMPG	75	QAHNVMRGRMPG	11	HSHYFIRGRIPG	2
QAHDVMRGCMPG	74	QPHDVMRGRMPG	11	QAHDVMRGRFPG	2
QEHDVMRGRMPG	74	QAHDVIRGRMPG	11	QAPDVMLGRMPG	2

QAHDVMRVRMPG	73	QAHEVMRGRMPG	11	QAHDVMRGRMPG	2
QAHDVVRGRMPG	73	QAHDVMRGRMPG	10	QAHDVMRGGMPG	2
QAHDVMRGRMPG	72	QAHDVMRRRMPG	10	QAHDVRRGRMQG	2
QAHDVMRGRMPG	72	QAHDVRRGRMPG	9	KAHDVMRGSMPG	2
QAHDVMRGRTPG	72	QAHDVMRGVCR	8	QAHDVCGGVCR	2
QTHDVMRGRMPG	71	QAHDVMRGPMPG	8	QAHYVMRGRMPG	2
QAHDVMRGRMPG	70	QAHDLMRGRMPG	8	QAHEGMRGRMPG	2
QAHDVKRGRMPG	67	QAHHVMRGRMPG	8	QAHAVMRGRMPG	2
QAHDVMRGRMPG	67	HAHDVMRGRMPG	6	QGHDVMGGRMPG	2
QAHDVMRERMPG	66	QAHDVMRGRMPG	6	QEHDVMRGRMPG	2
QAHDVMRRRMPG	65	QAHDVMPGRMPG	6	QAPAVMRGRMPG	2
QAHDVMRGRMPG	63	QAHDGMRGRMPG	6	*ADDVMRGRMPG	1
QAHDVIRGRMPG	63	QEHDVMRGRMPG	6	QAHDVMRGRMRG	1
HAHDVMRGRMPG	62	QAHDVMRGRMPG	5	HAHYVIRGRMPG	1
QAYDVMRGRMPG	62	QAHDVMRGRMPG	5	QADDVMRGRMAG	1
QSHDVMRGRMPG	62	PAPDVMRGRMPG	5	QAEDVMRGRMPG	1
QAHDVMRGRMPG	61	QAHDGMRGRMQG	5	RAQDVMRGRMPG	1
QAHDVMRGRMPG	60	QAHDVMRGRMPG	5	QAHYVMRGRMPG	1
QAHDVMRGRMP	60	QAHDVCGGVCR	4	QAHDVMPGRMPG	1
QAHDVMRGRMPG	59	HAHYVIRGRMPG	4	QTPDVIRGRMPG	1
QAHDVMRGRMPG	58	KAHDVMRGRMPG	4	QAHGVERGRGRG	1
QAQDVMRGRMPG	58	PAHDVMRGRLPG	4	QAHDVMRGANAG	1
QAHDVMRGRVPG	56	QAHYVMRGRIPG	4	QAHGVRRGRMPG	1

Of these 192 most frequent samples, 92.71% were 36bp as expected, 5.73% were 35bp, 1.04% were 37bp and 0.52% were 34bp, reflecting a combination of natural phage variation and possible sequencing errors. The 36bp samples were uploaded to MEME Suite for motif analysis [260]. There is a strongly conserved 12-residue motif observed (Figure 6.23). This shares the glutamine (Q) residue followed by basic and non-polar domains seen in the original small-sample analysis (Figure 6.19, Positions 10-13), but also offers a much more consistent motif than seen previously. The differences in sequence demonstrate both the importance of large-scale motif analysis and the potential for a higher-affinity binding peptide to be produced in the future for assay purposes.

It must be noted that latter half of the sequence shown in Figure 6.23 bears considerable similarity to the p16 motif posited as the binding site for the antibody used

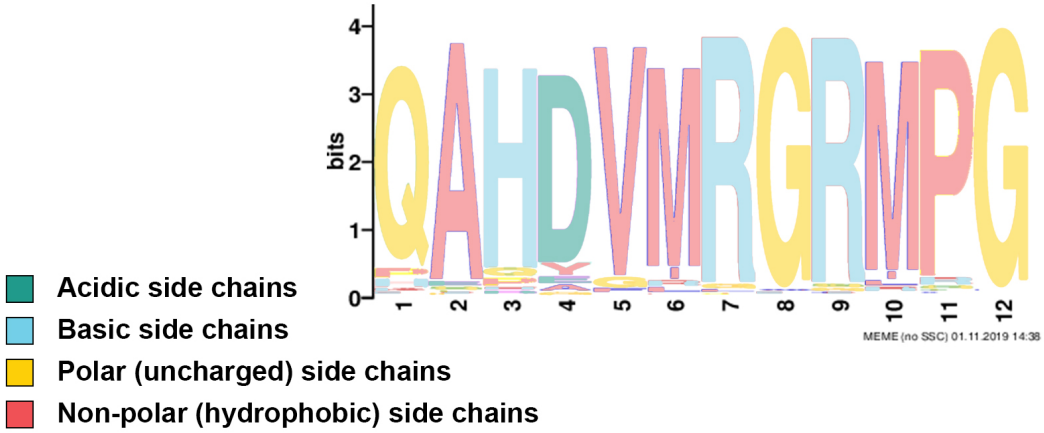


FIGURE 6.23: The 12-residue peptide motif produced as a result of analysing the 1,000,000 reads of 192 pooled samples, as analysed in MEME Suite.

(WGRLPV). It is possible that the peptides have preferentially bound to the antibody, although the similarity of sequence with p16-only binders would suggest otherwise. In either case the results are useful to the production of a paper-based assay, owing to the need for a control peptide. Further tests to confirm binding specificity are required.

6.7 SERS activity of paper-based devices

6.7.1 Choosing a substrate architecture

Initial investigations into the most appropriate vehicle for a paper-based assay were made through the exploration of constructed or commercially available dipstick assays and wax-printed 'dot blots', as described in Sections 2.6 and 2.7 of Chapter 2 respectively. Basic preliminary tests showed that paper assays were capable of producing an easily discernible Raman signal in the absence of a collimating pinhole (Figure 6.24).

A wax-printed nitrocellulose 96-well assay was prepared with a two-fold serial dilution ladder of purified p16 in PBS with 5µl added to each 'well' and left to dry at ambient temperature. Different absorption times and wetting effects were observed for the different concentrations of protein. The dry assay was then affixed to blotting paper using a spray adhesive (3M Photo Mount) before 5µl SERS-active conjugate (5µM BPE) was added to each well and left to flow through. The completed assay was interrogated using a Renishaw inVia upright microscope at 5x magnification with

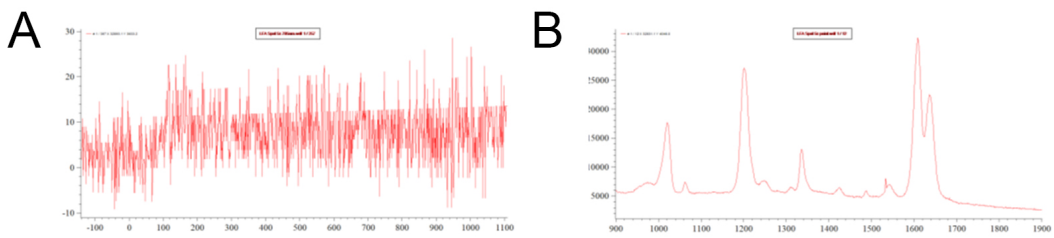


FIGURE 6.24: Raman spectra of initial dipstick tests, taken at 785nm with 1.0s single acquisition of the colorimetric spot. A) Spectral signature with pinhole. B) Spectral signature without pinhole, showing characteristic BPE peaks.

633nm laser. The brightfield microscopy and Raman peak intensity data can be seen in Figure 6.25.

The blotchy effect in Figure 6.25A is clearly the result of interaction with the droplets of spray adhesive, but there are also artefacts from the wax printing process. Colorimetric values were sampled in triplicate from non-artefact areas and can be seen in Figure 6.25B. Higher values correspond to colours closer to white in these 8-bit images, with each colour channel capable of displaying from 0 to 255 . Thus there is an inverse correlation with p16 concentration. The trend is extremely weak and could not form the basis of a quantitative test.

For the Raman interrogation, a map was generated using the process described in Section 2.8, Chapter 2. Principal Component Analysis (PCA) was carried out to isolate the chief source of variance - the SERS conjugate signal - from weak background intensities from the dyed wax or nitrocellulose membrane. Representative spectra from the maps were matched to the SERS signal principal component for each concentration. The intensities of the 1202cm^{-1} BPE peak are shown in Figure 6.25C. Again, there is a weak positive correlation.

It is likely that, due to the extremely shallow flow column of each well (101.6µm), small variations in adhesion to blotting paper generate significant contributions to the wicking properties and therefore Raman signal intensity of each well. Even on the basis of significant improvement to the adhesion method, variation in flow rate through the composite layers of the device also raises questions as to the quantitative robustness of such an assay.

Efforts were then focused on the use of dipstick assays, which overcome the Raman sampling issue by virtue of the lateral rather than depth-wise nature of the binding interaction.

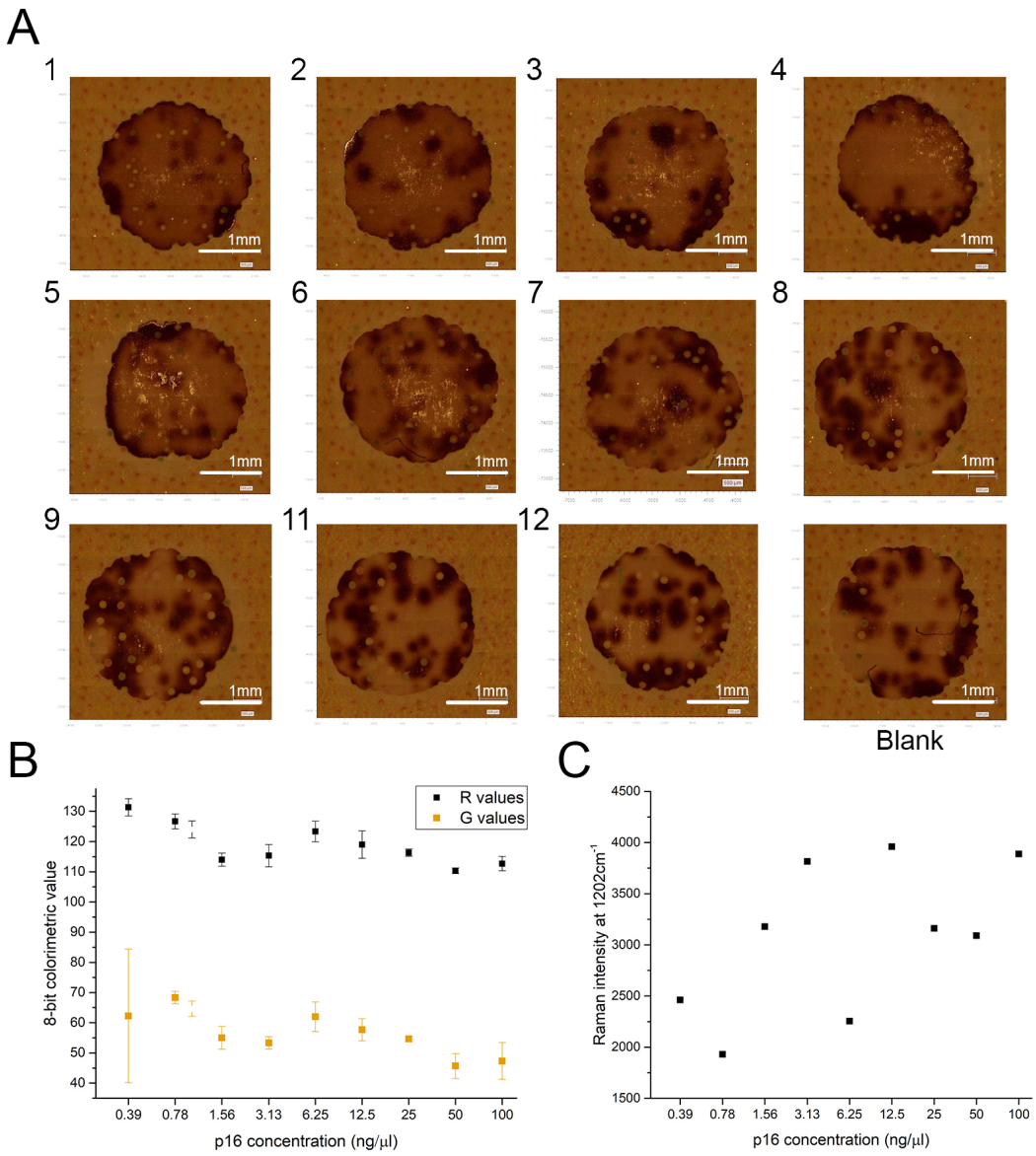


FIGURE 6.25: Conjugate response to purified p16 in a wax-printed assay format. A) Brightfield images of the spots, starting at 100ng/µl at position 1 in twofold serial dilution through to position 12. Glue aggregation (dark spot) and wax resist (light spot) artefacts are clearly visible. B) Direct colorimetric values from the brightfield red (R) and green (G) channels, averaged from analysis in triplicate. C) Single acquisition SERS response of p16 concentrations.

6.7.2 Raman response of assembled dipsticks

Dipsticks were assembled in house as described in Section 2.6.1, Chapter 2 and 1µl serially diluted purified p16 spotted onto the membrane as before, but starting at 100ng/µl. The assays were left to dry completely before applying 20µl conjugates mixed 1:1 with 1x conjugate buffer + 1% BSA (Expedeon). Owing to slow flowthrough another 20µl conjugate buffer was added to reduce the likelihood of high background. This was still a clear issue for the assembled assays. Nevertheless, the protein-bound areas were spot-sampled with a Raman spectrometer as described in Section 2.8 and scanned for colorimetric analysis as in Section 2.6.2.

For the colorimetric assessment, the scanned dipstick images were converted to greyscale to reduce dimensionality and then colour-sampled using a 5x5 pixel average in the spot centre and base of the strip, taken as a background value. Both were subtracted from 255 (8-bit absolute white) to give a scale which varies proportionally with colour saturation. For Raman analysis, single acquisitions were taken in the same manner using 10% laser power and a 1.0s acquisition. The results are summarised in Figure 6.26.

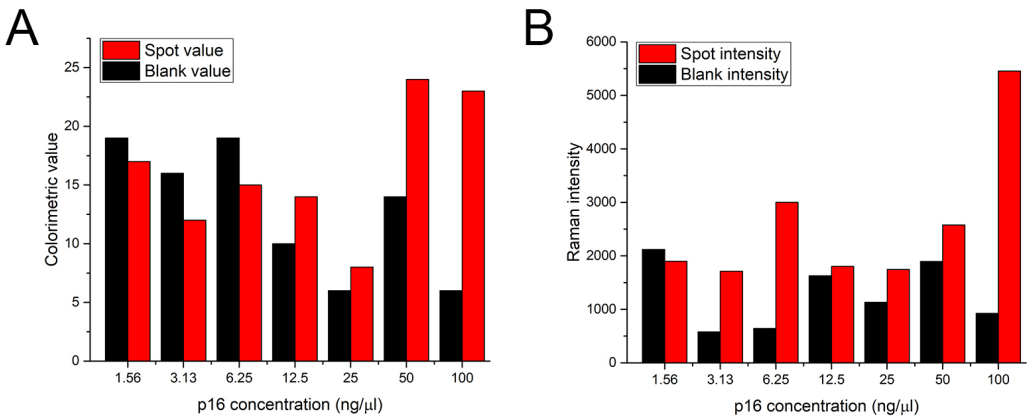


FIGURE 6.26: Conjugate response in assembled dipsticks. A) Inverted colorimetric values of the p16 binding region (spot value) and blank region (blank value). B) Single acquisition SERS response of p16 concentrations and corresponding blank region.

The colorimetric response in Figure 6.26A demonstrates no discernible trend, and in some cases is lower than the background signal for the strip. The Raman intensity results in Figure 6.26B also suffer from high background, but with a possible trend emerging for 25-100ng/µl p16 concentration. The primary issue for the assembled

dipsticks, demonstrated by high background and signal noise with both analysis techniques, is the variation in flow characteristics caused by the assembly itself. It was observed that the cellulose wicking pad in particular varied in terms of the thickness and quality of adhesion. If a paper-based assay were to be produced for clinical testing this would be overcome through automated manufacturing, and so the commercial dipsticks used previously for conjugation testing were adopted to remove this source of error.

6.7.3 Detection limits of commercial dipsticks

The conditions used in Section 6.7.2 were repeated with commercial dipstick assays optimised for use with gold antibody conjugates ('Conjugate Check and Go!', Expedeon) but with a lower starting concentration of 50ng/µl. Each concentration (50.00, 25.00, 12.50 and 6.25ng/µl) was run in duplicate. The completed assays were scanned for colorimetric analysis as before. Raman point acquisitions were also conducted with the same settings. Prior to spot sampling the SERS intensity, rapid Raman maps were acquired to ensure spectral consistency across the protein-bound area. This was carried out using StreamlineHR on Wire 4.4, with 5x magnification, 1.00s acquisition, 10% laser power, no pinhole, high gain and fast raster. All maps were 56x55 sampling pixels, with 20 and 12.5µm steps in X and Y respectively. The results, shown in Figure 6.27, were interpreted using differential classical least squares analysis (DCLS) with the bulk conjugate SERS signal as a reference component. The map colours were scaled to include the central 5-95% of spectral intensities.

A direct comparison of SERS response from the spots and associated background can be seen in Figure 6.28, which validates what was seen colorimetrically; the background signal is much lower for the commercial strips. The Raman and colorimetric results are seen in Figure 6.29. As seen in Section 6.7.2, the colorimetric results were 'inverted' such that higher values correspond to a more intense signal. For single point acquisitions an average of three measurements was taken for each sample, which was averaged again between the samples in duplicate. The spectral bands in Figure 6.29A indicate a single standard deviation on either side of the mean spectrum. The 1203cm^{-1} peak of the BPE spectrum was used for analysis and compared against the average blank SERS signal as seen in Figure 6.29B. The reference values do not intersect with the background signal, indicating that even vanishingly small concentrations of p16 could legitimately be compared to the reference curve.

For a more conservative estimate, the limit of detection (LoD) was calculated, given by

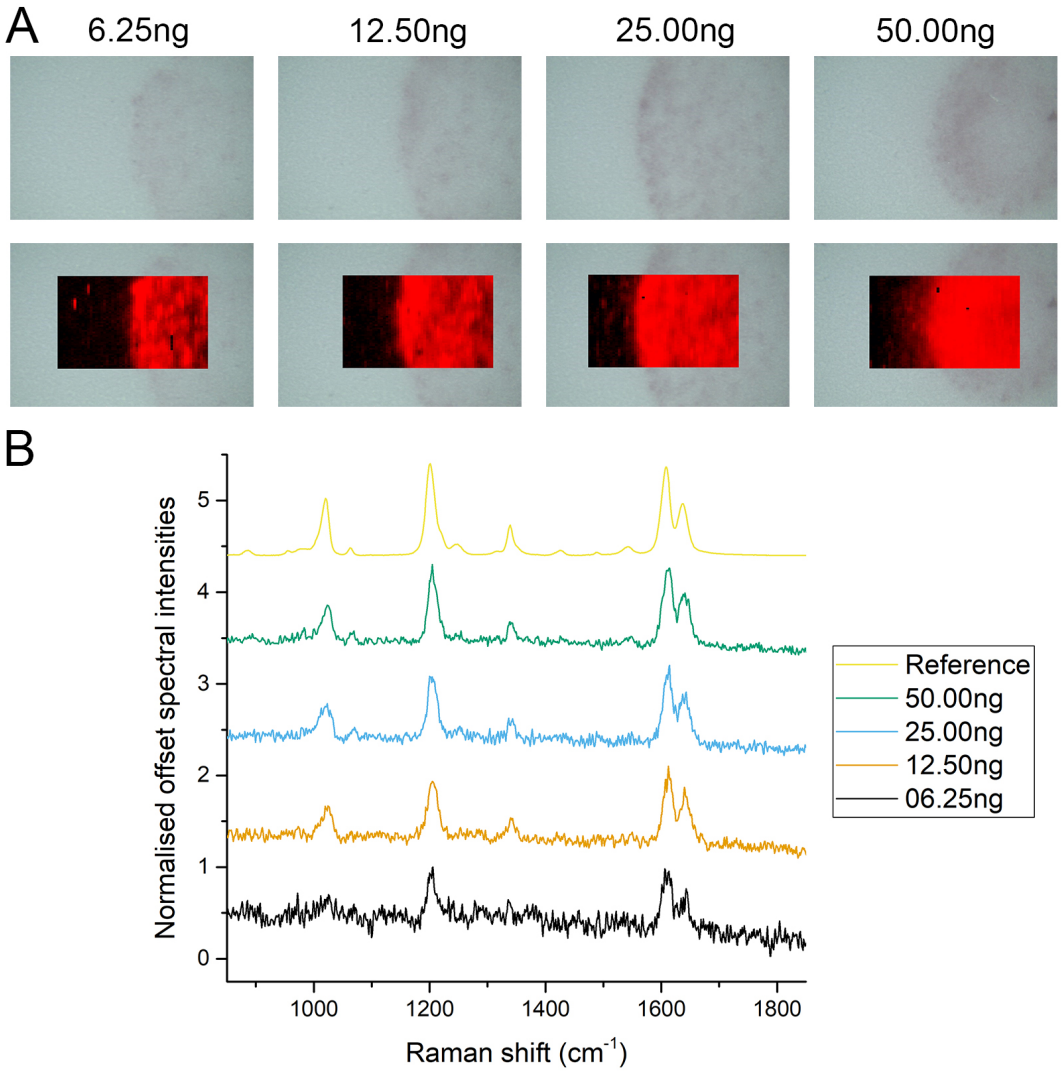


FIGURE 6.27: Raman maps of commercial dipsticks in the protein binding region. A) False colour Raman maps of colorimetric spot, created using a Classical Least Squares (CLS) fit of the spectral data using the conjugate bulk SERS signal as the reference component. B) Representative spectra from each level of p16 content and bulk conjugate reference.

$$LoD = LoB + x\sigma_{samples} \tag{6.1}$$

where x is dependent on the confidence interval chosen and the limit of blank (LoB) is given by LoB = $\mu_{blank} + x\sigma_{blank}$. In terms of Raman intensity for the 1203cm^{-1} peak, the values are $\mu_{blank} = 71.39$ and $\sigma_{blank} = 32.89$. If calculated as the average of sample standard deviations, $\sigma_{samples} = 100.33$. This gives LoB and LoD values as

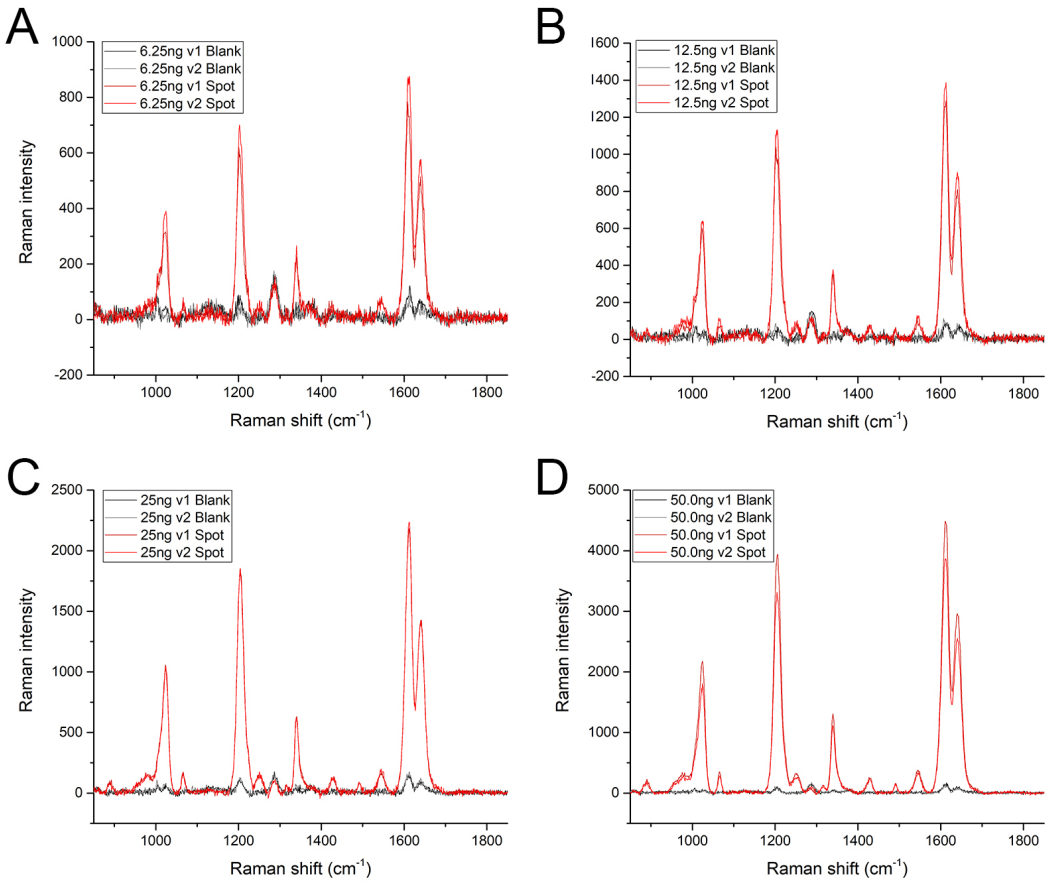


FIGURE 6.28: SERS spectra of commercial dipsticks at increasing concentrations of p16 (1µl volume used), compared directly to blank region response from each completed assay as an assessment of signal-to-noise ratio. A) 6.25ng. B) 12.50ng. C) 25.00ng. D) 50.00ng.

summarised in Table 6.9. In terms of absolute protein content, which represents the most suitable unit of measurement on the basis of the assay's capture mechanism, this is of a sufficient level of protein detection for wildtype SiHa cells (3-6ng) to a 99% confidence level. Further tests would be required to indicate whether this would be sufficient for detection in heterogeneous cell populations such as clinical samples.

The colorimetric data can be seen in Figure 6.29D. Here, the blank values are all zero giving non-existent background, but the reference curve also reduces to zero at 4.55ng p16 content. As such, this value forms the limit of detection for colorimetric analysis, showing that SERS enhancement of the paper-based assay offers higher sensitivity than colorimetric analysis . Furthermore, at this level of detection the wildtype

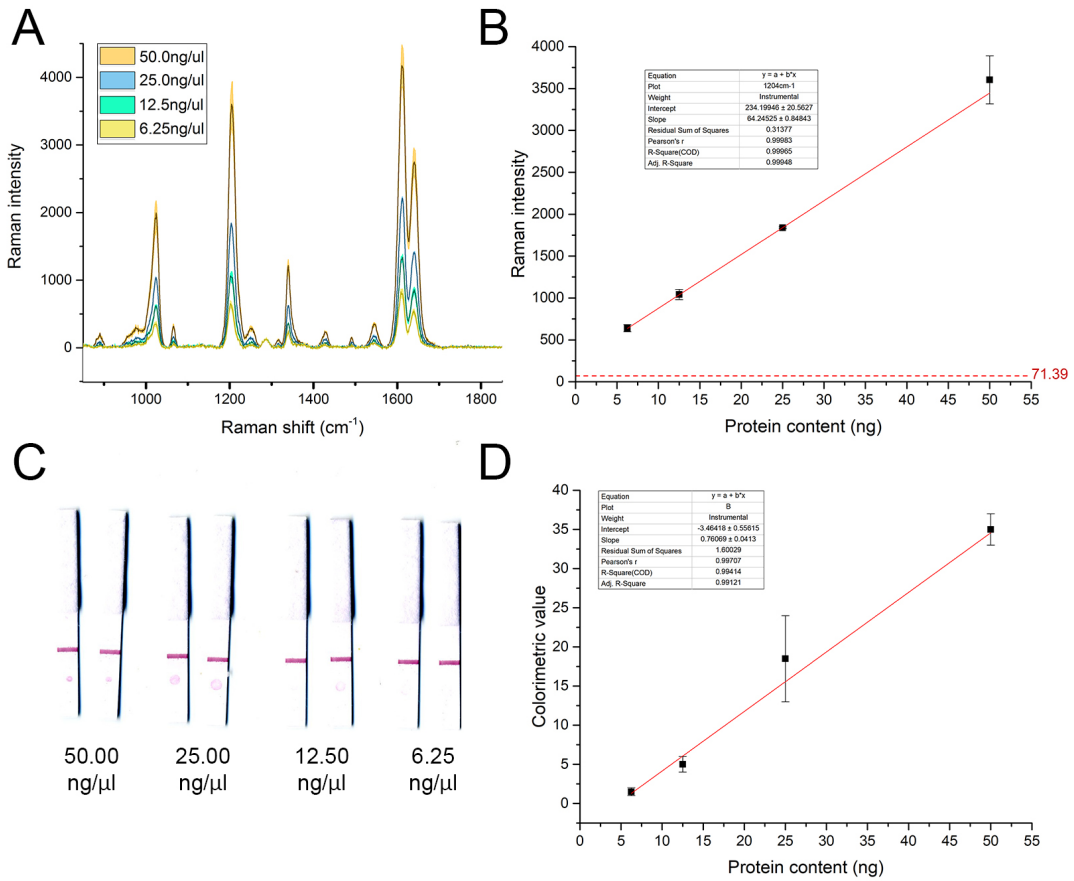


FIGURE 6.29: SERS and colorimetric responses of the commercial assays, tested as sample duplicates. A) The average spectral response as recorded from three single acquisitions of each assay, averaged again over sample duplicate. B) Peak intensity at 1203cm^{-1} characteristic peak of BPE with average assay blank SERS value (dotted line). C) Scanned images of assays used for colorimetric analysis. D) Colorimetric analysis, taken from a 5x5 average in the spot centre.

expression levels of p16 in SiHa would not be detected, showing that SERS enhancement is necessary for a quantifiable paper-based test.

Confidence (%)	x	LoB (Raman)	LoD (Raman)	Protein (ng)
90%	1.645	125.49	290.54	0.876
95%	1.960	135.85	332.51	1.530
98%	2.326	147.89	381.26	2.289
99%	2.576	156.11	414.57	2.807

TABLE 6.9: The limits of blank and detection for SERS paper assays in terms of Raman intensity, along with the equivalent protein content determined by linear extrapolation.

6.8 Conclusions

In this Chapter, a paper-based diagnostic device for p16 expression in fixed cell samples has been proposed and characterised in conjunction with the SERS-active antibody conjugate developed in Chapter 4.

Capture layers consisting of biopanned phage-displayed peptides were discovered, analysed and synthesised ready for addition to a streptavidin-printed membrane, with opportunity for high affinity motifs to be produced following residue analysis. Eight of the 24 panned antigen sequences were found to be significant in terms of binding affinity compared to secondary-only control. Of the top four, one sequence (ASHYHINYRKWT) provided easily discernible detection of purified p16 to 62.5ng/ml in a sandwich format, as well as significant differentiation between cell line models using native and transfected expression of target.

Next-generation sequencing of a larger phage population served both to validate initial results and search for other high-afffinity motifs. The NGS analysis yielded 192 unique motifs of frequency of 1 to 860,000 which produced, overall, a highly conserved 12-residue motif which can form the basis of further binding sensitivity experiments.

Having laid foundations for capture layer development, the detection mechanism and assay substrate were considered in Section 6.7. Commercial dipstick assays, produced by Expedeon, were identified as the basis for further assay development owing to their low background and consistent flow characteristics. Comparisons of colorimetric and SERS analysis of conjugate binding to p16 on the chosen substrate indicated detection limits of 4.55ng and 0.88ng to 90% confidence respectively, showing SERS to be the more sensitive imaging mechanism. In addition, SERS was able to detect p16 content at biologically relevant levels not found with colorimetric analysis.

Significant progress towards a feasible clinical assay has been made, but the assay clearly requires further development. Orbital raster scanning could be used to sample the whole spot, rather than mapping or spot sampling. Suggestions for further work in

the long term are considered in Chapter 7. In the immediate future, the peptide motifs for the test and control lines should be optimised for target affinity by testing a variety of motif candidates according to the analysis performed in Section 6.4. Following this, the biotin-tagged motifs should be bound to streptavidin-printed nitrocellulose dipsticks so that whole system testing can be carried out in earnest using cell models and eventually, clinical samples.

Through combination with a handheld Raman spectrometer and following the steps outlined above, this assay has the potential to be developed in keeping with ASSURED criteria for diagnostic use with smear samples.

Chapter 7

Conclusion and suggestions for further work

In this thesis, a broad range of enhanced Raman spectroscopic techniques have been explored and considered in a clinical context before a paper-based device for p16 quantification was judged to be the most feasible method and selected for further development.

In the global imaging approach, a previously developed wavelength modulated spectrometer was used to automatically classify fluorescence-free spectra of PreservCyt-fixed cultured cells in an analogue for cervical smears. PCA-based modulated spectral classification provides clear benefits over standard spontaneous classification. For comparison of HPV status in cervical cancer lines C33A (HPV negative) and SiHa (HPV16 positive), sensitivity and specificity were 88.0% and 90.5% respectively for standard classification, rising to 100.0% and 95.8% for modulated classification. For comparison of transformation status in normal (NHEK), spontaneously immortalised (HaCaT), cervical carcinoma in situ (SiHa) and metastasised (CaSki) cells, the standard sensitivity and specificity were 84.8% and 85.1% respectively, again rising to 95.1% and 96.2% for modulated classification.

Intracellular sampling location studies indicated that fixed harvested cells were less sensitive to sampling location than live, adhered cultures, which may have practical benefits in terms of sampling tolerances. Nonetheless, the sensitivity of the technique and the discriminatory nature of PCA means that statistically significant numbers of clinical samples would require interrogation to guarantee disease-specific spectral classification.

As discussed in Chapter 3, the sampling speed and acquisition method of this technique are the major barriers to sufficient throughput and therefore clinical utility. Although the introduction of primary HPV testing will significantly reduce the number of samples requiring cytologic analysis and the automated system would not

require office-hours supervision as in current practice, a single smear would still have to be processed in approximately 162min or 1s per cell in order to 'break even', based on HPV incidence [261].

The possibility of an optical coherence tomography or digital holographic microscopy map co-registered with the current point-sampling system has been posited as a way of removing the need for raster scanning empty areas of the slide. Coupling with existing digital image guidance systems would achieve the same purpose. There is also the possibility of a flow cell-style setup, with cells held in place with optical tweezers during interrogation. Again, significant improvements to acquisition speed would be required. Finally, light-sheet Raman spectroscopy would allow for multiple spectra to be acquired simultaneously and deconvoluted post-capture. Potential models for such a system include selective plane illumination microscopy (SPIM) with Fourier transform analysis [262] or digital scanned lightsheet microscopy (DHLM) with an interferometric tunable filter [263]. Certainly, a global method in this form would offer enormous benefit not just for cervical smears, but for any pathology where cell sampling is useful and feasible.

A functional SERS-active nanoparticle conjugate to p16 was then developed for consideration of targeted approaches. The three most common methods of conjugation have been assessed and optimised, including EDC/NHS covalent, passive and Protein G' bioconjugation. The final Protein G' conjugate described can be produced in 2.5 hours at room temperature at a material cost of £2.32 per 1ml (unconcentrated), demonstrating baselined Raman intensity of 11530 ± 2650 counts without concentration and a clear colorimetric and Raman response to 50ng purified p16 in a paper-based assay. A workflow was also developed for strategic approaches to general antibody-nanoparticle conjugation in the case where maintaining binding functionality proves difficult.

Two targeted approaches to Raman spectroscopic classification of cervical smears were broadly explored using the developed conjugate, with a third considering a hybrid global-targeted approach using alkyne-labelled antibodies with stimulated Raman spectroscopy.

In the more established SERS mapping framework, a dummy conjugate was developed for ratiometric analysis of whole cell p16 expression which demonstrated a calibrated system capable of translating ratiometric conjugate intensity into the percentage volume of target. This allows for binding specificity of the initial 1:1 mix to be considered, but this requires further validation in terms of expected expression levels in clinical samples. In the cell analogues tested, there was a 2.2-fold Raman intensity increase for transfected versus endogenously expressing SiHa cells. Preliminary

method optimisation indicated that PreservCyt fixing leads to lower non-specific accumulation of conjugates within cells.

As with global point sampling, this approach still suffers from relatively cumbersome acquisition, although it does offer improvements on sampling speed in comparison with the wavelength modulated approach. SERS mapping could be combined with topographical or image guidance methods as described above for increased acquisition rate per sample, or by considering overall signal from a pooled cell pellet. However, this offers no real advantage over fluorescent labelling unless marker multiplexing levels were high. In the multiplexing scenario the practicalities of cost in a clinical setting would become an issue and is irrelevant for screening which incorporates a robust and well-established marker.

The SERS-SRS method, in which a SERS-active nanoparticle conjugate is used as a labelling moiety in the global SRS method, was insufficiently developed as a technique to justify further work at this time, especially in a fixed cell context which does not require the use of nanoparticle stealth properties or multiplexing.

Alkyne labelling was then taken up for further study, owing to the availability of strongly Raman active molecules for SRS labelling without the need for considerations of nanoparticle stability. A method was optimised for succinimide-conjugated antibody amines using a fluorescent analogue, meaning that alkyne labelling can be continued in earnest in the future. This global-targeted hybrid method, which allows for simultaneous consideration of specific markers with global morphological information from key vibrational modes, offers the prospect of a genuine advance on established fluorescent imaging which could be expanded to histological studies and biomarker discovery within the molecular pathology framework. It is suggested that upon successful production of an alkyne-labelled antibody, a comparison of global only, target only, and global-target hybrid approaches are compared directly with each other and traditional cytology to quantify the extent of classification power. The global method could use significant vibrational modes discovered in Chapter 3.

The targeted approaches explored clearly offer benefits in a diagnostic context, but these benefits are not suited to cervical screening. After consideration of the strengths and weaknesses of the trialled targeted approaches, a different paradigm was adopted which used the SERS-active conjugates in a paper-based assay setup similar to that used for optimisation work. This capitalised on the advantages of the targeted SERS approach in the cervical cancer context, allowing for sensitive, point of care quantification of a relevant marker which has the potential to transform screening practices in a variety of settings, conforming with ASSURED criteria.

Phage-displayed peptide biopanning was used to discover potential binding motifs for p16 protein and antibody. This was first carried out on a small scale, with one sequence providing easily discernible detection of purified p16 to 62.5ng/ml in a sandwich format, as well as significant differentiation between cell line models using native and transfected expression of target. Validation through larger scale next-generation sequencing produced a high-consensus alternative motif ready for binding validation tests. A commercial dipstick assay was used to test SERS activity of conjugates in a paper-based format. Comparisons of colorimetric and SERS analysis of conjugate binding to p16 on the chosen substrate indicated detection limits of 4.55ng and 0.88ng to 90% confidence respectively, showing SERS to be the more sensitive imaging mechanism in this instance. In addition, SERS was able to detect p16 content at biologically relevant levels not found with colorimetric analysis. As such, the assay components offer genuine functionality in a clinical context, but require further development.

Beyond immediate optimisation and characterisation of the discovered peptide motifs through alanine scanning, binding assays and the discovery of the antibody binding site through the same method, the identified components should be combined into a functional assay. First, biotinylated peptides should be printed onto nitrocellulose strips using streptavidin adhesion zones. It is possible that successful printing and binding could be tested using an adapted Western Blot method for the targets in question. Once the assay architecture itself has been validated, the substrate and conjugates should be tested with analytes in a variety of formats, including a spiked purified p16 concentration ladder, lysed cell models, and finally clinical samples. It is recommended that an industrial partnership with Expedeon should be formed in order to develop a scalable manufacturing process for the assays.

Further optimisation of the detection sensitivity is also recommended. Basic collection performance improvements such as improved lighting and analysis of colorimetric data, plus longer acquisition times for SERS measurements subject to substrate degradation, would give a more realistic understanding of the technical limits of each technique. Smaller nanoparticles (30nm) would yield both a higher colorimetric contrast and SERS amplification, but may also be subject to lower stability. In this case, optimisation of the excitation wavelength could be considered an alternative whilst retaining the current conjugate. The 40nm gold conjugates produced were designed for optimal SERS enhancement at 532nm excitation, but were found to give highly variable signals. This is posited to be due to excessive heat conversion at the LSPR peak, which may have degraded the surface molecules.

One of the key criteria for a robust assay in this context is the ability to threshold p16 values across a whole sample so that they can be classified with reasonable levels

of confidence. Although outcome of borderline samples could still be confirmed by a cytopathologist, the number of samples requiring this pathway would have to be relatively small for the test to significantly reduce cost and workload. To date studies of p16 stain or p16/Ki67 dual stain have been conducted with conventional cytology samples consisting of discrete cells, and so a whole-sample study of clinical samples or modelled analysis of existing stained cytology slides would be required, preferably in comparison to colorimetric methods to ascertain the relative benefit of SERS-based paper assays in relation to their additional complexity. In the case that p16 alone has insufficient resolving power, the assay could easily be adapted to include a test for Ki67 as discussed in Chapter 1 - indeed, this is recommended for any possible clinical introduction in a non-morphological format.

The diagnostic landscape and future avenues explored in this thesis are summarised in Figure 7.1. It is hoped that this research will provide a useful reference for all those working in diagnostic and Raman research contexts, as well as resources required to create a point-of-care, inexpensive test for p16 which could streamline and broaden access to cervical screening across the globe.

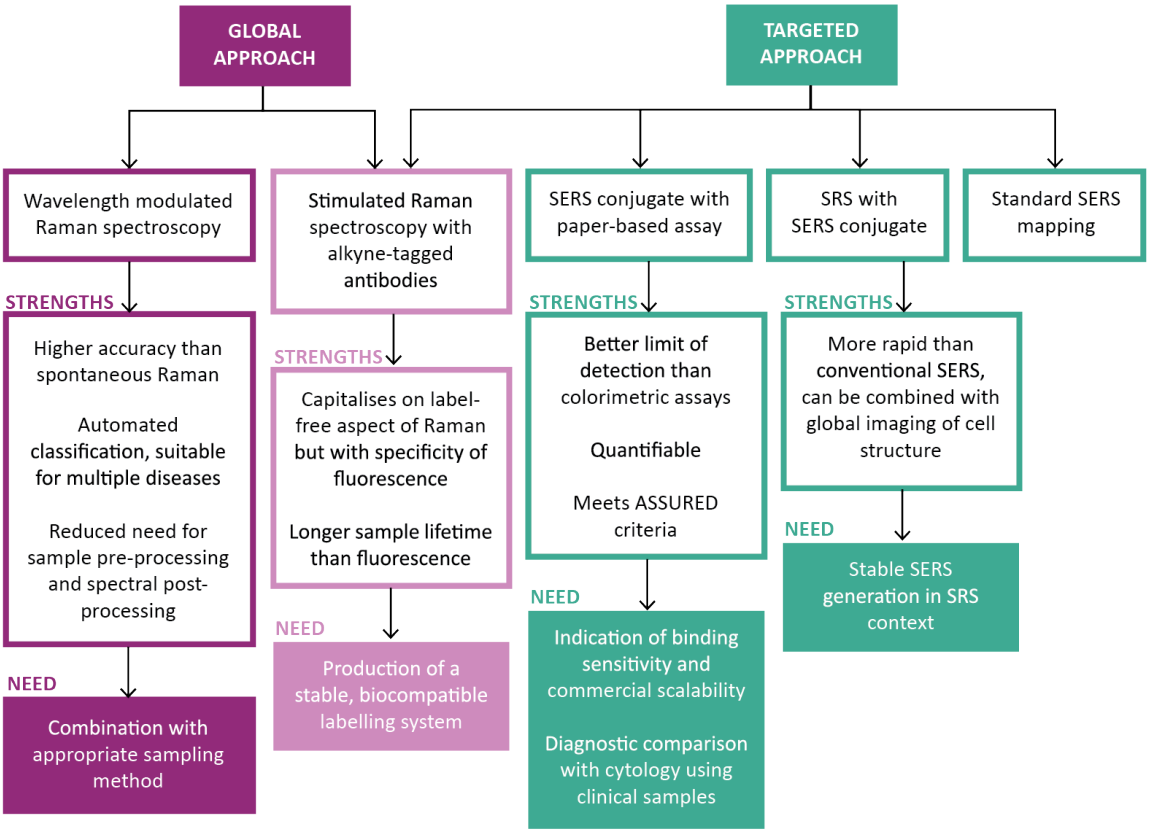


FIGURE 7.1: A summary of strengths and suggested directions for Raman spectroscopic studies in different contexts. Standard SERS mapping strengths are not listed as they were not considered to be relevant to cervical screening.

Chapter 8

Appendix A: Published research

This appendix contains a reproduction of the paper "Towards automated screening ", published in the Journal of Biophotonics online in Janurary 2018 and in print in April 2018.

Received: 11 September 2017 | Revised: 24 November 2017 | Accepted: 23 December 2017

DOI: 10.1002/jbio.201700244

Journal of BIOPHOTONICS

FULL ARTICLE

Towards automated cancer screening: Label-free classification of fixed cell samples using wavelength modulated Raman spectroscopy

Lana Woolford[1] | Mingzhou Chen[2] | Kishan Dholakia[2] | C. Simon Herrington[1]*

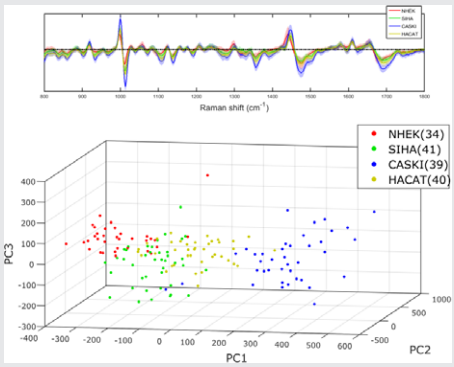
[1]Edinburgh Cancer Research Centre, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[2]SUPA, School of Physics and Astronomy, University of St. Andrews, Fife, UK

***Correspondence**
C. Simon Herrington, Edinburgh Cancer Research Centre, Institute of Genetics and Molecular Medicine, University of Edinburgh, Crewe Road South, Edinburgh EH4 2XR, UK.
Email: simon.herrington@ed.ac.uk

Funding information
Engineering and Physical Sciences Research Council, Grant/Award number: EP/L016559/1; Medical Research Council, Grant/Award number: EP/L016559/1; Cancer Research UK, Grant/Award number: A18075

The ability to provide quantitative, objective and automated pathological analysis would provide enormous benefits for national cancer screening programmes, in terms of both resource reduction and improved patient well-being. The move towards molecular pathology through spectroscopic methods shows great promise, but has been restricted by spectral quality, acquisition times and lack of direct clinical application. In this paper, we present the application of wavelength modulated Raman spectroscopy for the automated label- and fluorescence-free classification of fixed squamous epithelial cells in suspension, such as those produced during a cervical smear test. Direct comparison with standard Raman spectroscopy shows marked improvement of sensitivity and specificity when considering both human papillomavirus (sensitivity +12.0%, specificity +5.3%) and transformation status (sensitivity +10.3%, specificity +11.1%). Studies on the impact of intracellular sampling location and storage effects suggest that wavelength modulated Raman spectroscopy is sufficiently robust to be used in fixed cell classification, but requires further investigations of potential sources of molecular variation in order to improve current clinical tools.



KEYWORDS
cancer, cervix, Raman spectroscopy, screening

1 | INTRODUCTION

The introduction of national screening programmes such as the Papanicoloau ('Pap') test and mammogram have achieved considerable success, and are estimated to prevent, for example, 4500 [1] and 1300 [2] deaths in the United Kingdom every year from cervical and breast cancer, respectively. However, these programmes require large, sustained financial input and human resources and are not completely accurate. The primary purpose of cervical screening is to detect abnormal epithelial cells that indicate the presence of cervical pre-cancer or cancer. In a cervical screen, cells are brushed from the cervix and stored in fixing solution as a liquid-based cytology sample. This 'smear' is transferred to a microscope slide, stained and viewed by a cytopathologist, who classifies it as normal, borderline, low

or high grade dyskaryosis, glandular neoplasia or unsatisfactory [3]. However, interpretation of cell morphology is hugely complex and therefore, open to subjectivity and false diagnosis, as small numbers of abnormal epithelial cells must be identified in a background of numerous normal epithelial and inflammatory cells. In 2008, clinical trial comparing reviewed with initial pre-treatment histologic diagnosis, the probability of unnecessary treatment was found to be 8% to 27% (dependent on initial cytologic findings) [4], indicating that many women may be being referred to colposcopy unnecessarily. The use of biomarker screening has been posited as a potential method for improving diagnostic accuracy, but identifying targets relevant to genetically heterogeneous populations exposed to a variety of environmental factors while maintaining high-positive predictive value is often unfeasible [5]. Thus, a technique which provides the benefits of quantitative molecular pathology but with a broader scope and automated classification could vastly improve screening outcomes.

One technique with the potential to deliver these characteristics is Raman spectroscopy. Raman spectroscopy is a label-free technique based on the inelastic scattering of light. It can be used to identify the molecular vibrational bonds present in a sample. Changes in disease state are accompanied by alterations in molecular pathology, which are reflected in spectral data. Raman spectroscopy in its various forms is emerging as a key contender for analysis of cancer status [6], including colorectal [7], lung [8] and cervical cancer [9–11]. However, the likelihood of clinical uptake is highly dependent upon the ability to contend with the practicality as well as accuracy of diagnostic gold standards such as histopathology. As such, a new classification method would need to be incorporated into current systems for trials to be conducted in the first instance.

In this paper, we demonstrate the use of wavelength modulated Raman spectroscopy (WMRS) [12] as a tool for the automated classification of unlabeled fixed cells in suspension, reflecting its potential utility in diagnosing clinical liquid-based cytology samples. One of the major technical barriers to the introduction of Raman spectroscopy as a clinical tool is the long acquisition time associated with the rare event of Raman scattering of photons. The acquisition of high-quality spectra is further compounded by the presence of autofluorescence in samples, which is a particular problem for biological interrogation and can often lead to the loss of smaller Raman peaks obscured by a fluorescent background. Several approaches for reducing the background have been explored, including post-acquisition baseline correction, nonlinear optical methods and time-, frequency- and wavelength-domain methods [13].

WMRS is a wavelength-domain approach. During measurement, the laser wavelength is modulated throughout the acquisition process using a tunable laser [14]. This causes the Raman spectral peaks to shift fractionally when measured in terms of scattered wavelength. When this shift is measured across several wavelengths, a 'differential spectrum' can be generated which demonstrates the true Raman peaks in wavenumber shift from the incident wavelength. Wavelength modulation does not affect the position of autofluorescence from the sample and substrate, allowing the background to be automatically removed [12]. Methods for post-processing wavelength modulated spectra in order to account for wavelength shift and remove this static fluorescence background have been tested, including standard excitation Raman difference spectroscopy, SD, Fourier filtering, least-squares fit and principal component analysis (PCA), with PCA providing the best signal-to-noise ratio [15]. The differential Raman spectrum is formed from the first principal component (PC) of the modulated data set, corresponding to the component which captures the greatest variance in the data.

This method has been used to investigate single living Chinese Hamster Ovary cells (CHO-K1) [12] as well as T lymphocytes ($CD4^+$, $CD8^+$), killer T cells ($CD56^+$) and dendritic cells ($CD303^+$, $CD1c^+$) [16]. It has also been shown to provide high specificity and sensitivity for the characterisation of human urothelial (SV-HUC-1) and bladder cancer (MGH) cell lines cultured normally [14] or exposed to urine [17], which represents another step towards an automated optical platform for cancer screening. The aim of this study is to assess the practical potential of WMRS as a method for the automated classification of fixed squamous epithelial cells, modelling cervical cancer diagnosis.

2 | MATERIALS AND METHODS

2.1 | Cell sample preparation

A total of 4 established cell lines, and primary cells, from squamous epithelial tissue were used for this project, reflecting the intention to investigate WMRS as a method of classifying cervical smear samples. CaSki, SiHa, C33A (all squamous cell carcinoma, adult human cervix), HaCaT (immortalised squamous epithelium, adult human skin) and primary normal human epidermal keratinocytes ('NHEKs', normal squamous epithelium, neonatal human skin) were cultured to 80% to 90% confluency. These were harvested by trypsinisation, washed in triplicate by centrifugation and resuspension in sterilised phosphate buffered saline (PBS), and fixed in the clinical methanol-based fixative PreservCyt. Samples were stored in PreservCyt at −20°C until use. All centrifugation steps were executed at 400. for 5 minutes at 4°C. Prior to interrogation, fixed cell samples were washed in triplicate by centrifugation and resuspension in sterilised PBS in order to remove traces of fixing solution, as methanol produces strong Raman signatures. Twenty microliter of PBS cell suspension was then transferred to a quartz cavity slide (Figure 1), carefully sealed with a 150 μm quartz

coverslip and inverted. The quartz assemblies were rested for at least 20 minutes prior to interrogation in order to promote cell adherence to the quartz coverslip, which was placed face down on the inverted microscope assembly for measurement. Details of the origin and cell culture reagents used for each cell type are provided in Appendix S1, Table 1.

2.2 | Spectral acquisition

A confocal Raman microscope built and modified for acquisition of both standard and modulated spectra (Figure 2) was used to evaluate the fixed cell samples. Raman spectra were generated from individual cells inside the quartz cavity assembly using a tunable continuous wave Ti-Sapphire laser ('SolsTis', M Squared Laser Ltd, maximum power 1 W at 785 nm), delivering 150 mW power in a 1 μm spot in the sample plane through a 40× oil immersion objective (Nikon, NA1.3). A translational stage allowed the laser to be focused on the nuclear region of each cell, with Raman scattered photons collected from a confocal collection volume 2.86 μm in depth and 2.63 μm in diameter about the laser spot. The signal was resolved with a Shamrock SR-303i spectrometer (400 lines/mm grating, Andor Technology) and collected with a CCD camera (Newton, Andor Technology) thermoelectrically cooled to −70°C. 'Standard' spontaneous and wavelength modulated Raman spectra were acquired using Solis (s) Software (Andor Technology) along with a brightfield microscope image (Imaging Source USB Camera) of each cell (see Supplementary Information S3). Standard acquisitions were conducted at 785 nm, using 5 accumulations of 2 seconds laser exposure for both the cell and adjacent background. The adjacent background, including any spectral contribution from the quartz slide, was removed from the raw cell spectrum for data analysis. In the modulated acquisitions, 5 accumulations of 2 seconds exposures were performed with 5 different cycles of excitation wavelengths over a total range of $\Delta\lambda = 1$ nm centred at 785 nm. No background acquisition was necessary here due to the inherent autofluorescence removal. It has been shown previously that no significant phototoxicity was experienced by the cells as a result of laser exposure at similar settings [16].

Reference spectra for laser calibration were acquired from 1 μm polystyrene beads in distilled water. These spectra were used to account for fluctuations in laser power, wavelength and spot location due to heating, fibre-optic coupling or mirror alignment, as well as post-acquisition spectral processing.

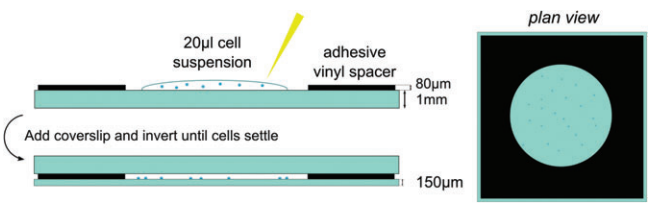


FIGURE 1 The cavity slide assembly used for all acquisitions. Slides were sterilised with ethanol, air-dried and polished in between measurements

2.3 | Spectral processing and data analysis

All spectra were considered in the fingerprint region of interest, 800 to 1800 cm^{-1}. Spectra were normalised according to the total area under each modulation set, which is linearly proportional to the laser power. Laser wavelength was also monitored by linearly interpolating the intensity and Raman shift of the characteristic 1001.4 cm^{-1} polystyrene peak according to the sample measurement time stamp and adjusting spectral position accordingly. Overall, this accounts for any laser power and wavelength fluctuations over time, thereby preventing the PCA script from grouping cells according to these parameters. Immediately prior to analysis, individual spectra were assessed and removed from the analysis if a high spectral background was apparent. Such artefacts arise from suboptimal focusing or cell movement during acquisition. All spectra were analysed using PCA, a commonly used multivariate spectral analysis technique [18]. PCA breaks spectra down into orthogonal spectral components PCs which describe the variation in the dataset. The first principal component describes the greatest variation, with higher components eventually describing noise [19]. The first 7 PCs were used in all data analysis conducted. The weighting of these PCs (the extent to which each PC represents a particular spectrum in the dataset) was used to generate PCA cluster plots, which may be used for diagnostic classification. Sensitivity and specificity values were generated through 'leave-one-out cross-validation (LOOCV)', also known as a 'jacknife' [18]. When performed for 2 datasets, the classifications can be quantitatively estimated as sensitivity and specificity values by assigning the each cell line as 'positive' or 'negative'. Statistically significant variations between spectra were also estimated for each cell line pair using student's . test at a significance level of . $< 10^{-12}$. This was conducted using the established 'ttest2' function in MATLAB, where the average spectral intensity at each wavenumber is compared, with consideration to spectral variation across the dataset (indicated in figures by coloured bands around the average). The PCA script used was written in MATLAB by the Biophotonics Group at the University of St Andrews.

3 | RESULTS

3.1 | Comparison of human papillomavirus status

A preliminary comparison of standard and wavelength modulated Raman spectroscopy as a classification tool was made using fixed C33A human papillomavirus (HPV

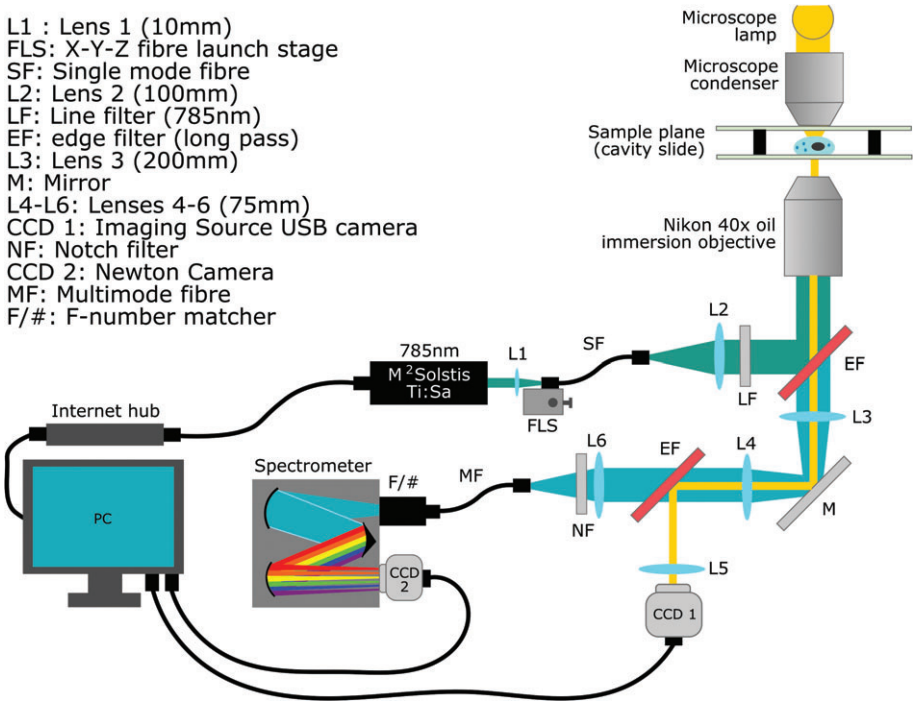


FIGURE 2 The optical diagram of the spectrometer designed for WMRS and standard Raman spectral acquisition. Brightfield microscope images are acquired separately on CCD1 and blocked from the spectrometer by secondary EF

negative) and SiHa (HPV16 positive) cells, kindly donated by the Scottish HPV Archive [20]. The samples were resuspended in PBS and prepared for analysis as described in Section 2.1. Initial analysis of the C33A cell line found that an acquisition time of 2 seconds over 5 accumulations yielded superior standard Raman spectra to a single acquisition of 10seconds. This is due to a comparative reduction in CCD camera noise for the shorter acquisitions. The brightfield cell images were morphologically similar, consisting of 10 μm cells which were highly rounded due to the trypsinisation process. The PCA plots for the standard and modulated Raman can be seen in Figure 3. LOOCV was applied to the datasets to yield sensitivity values of 88.0% and 100.0% for standard and modulated Raman, respectively, representing a considerable improvement of 12%. Specificities of 90.5% and 95.8% were found for standard and modulated Raman respectively, or a 5.3% improvement. It is clear from both PCA and LOOCV that modulated Raman spectroscopy provides a superior separation and classification of 2 cervical carcinoma cell lines with differing HPV status. The main spectral disparities occur at approximately 730, 760, 890, 920, 1130, 1250, 1380 and 1480 cm^{-1}.

3.2 | Comparison of different status in squamous epithelial cell lines

The investigation of standard and modulated Raman spectroscopy was then extended to 4 cell lines (1 normal, 1 immortalised and 2 transformed), which were chosen to represent the spectrum of abnormality typically found in the progression to cervical cancer. NHEKs are normal squamous epithelial cells which stop dividing when they reach a certain age. HaCaT cells [21] are a spontaneously immortalised epithelial line which can proliferate beyond this normal limit. These were used as a model for cervical precancer. SiHa [22] and CaSki [23] are cervical cancer cell lines from primary and metastatic tumours, respectively. These are both transformed phenotypes, consisting of cells capable of forming tumours. A total of 163 cells were interrogated using both standard and wavelength modulated Raman spectroscopy, consisting of 42 NHEK, 41 SiHa, 40 CaSki and 40 HaCaT cells. As before, the established cell lines in particular could not be discerned on a purely morphological basis, which furthers the case for biochemical classification as a complementary technique. It must be noted that the morphology of the primary NHEK cells varied significantly in terms of both shape and size compared to the established cell lines. Following removal of suboptimal spectra, Raman spectra and PCA plots were generated as seen in Figures 4 and 5, respectively. The ability of PCA to act as a classification algorithm was explored using cross-validation of the data. LOOCV was performed on all 4 cell lines for both datasets; the confusion matrices for standard and modulated PCA can be seen in Table 1 and 2, respectively, along with misclassification statistics. For all 3 established cell lines, the modulated spectra provide better discrimination. However, classification is slightly superior using standard Raman spectra for NHEKs.

In order to gain further understanding of both the efficacy of the test and the major vibrational signals contributing to the analysis, each possible pair combination of cell lines was analysed using the PCA script. The average sensitivity was 84.8% ± 9.0. and 95.1% ± 5.6., and specificity was 85.1% ± 10.7. and 96.2% ± 4.6., for standard and modulated PCA respectively. This demonstrates that, for both metrics, wavelength modulated Raman spectroscopy

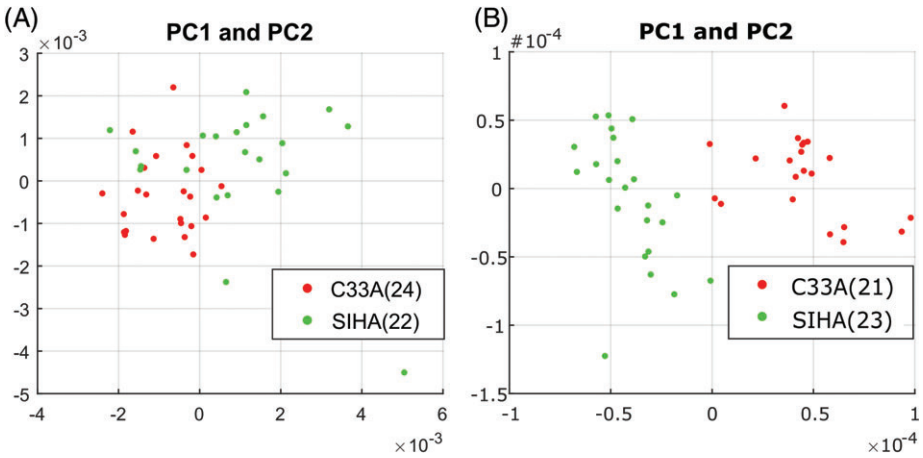


FIGURE 3 Standard (A) and modulated (B) spectral datapoints represented as PC1/2 weightings for C33A (HPV negative carcinoma) and SiHa (HPV-16 carcinoma) cells. The number of cells used for each comparison are indicated in the legend

provides superior discrimination (+10.3% sensitivity, +11.1% specificity) as well as lower variation between cell types. The generated spectra were also marked with significance bands as decribed in Section 2.3. Spectra with marked bands are shown for NHEK-Siha, NHEK-CaSki and NHEK-HaCaT are shown in Figures 6 and 7. The . test statistics used did not identify any bands for the combination HaCaT-SiHa, for either standard or modulated Raman. The lack of statistically significant spectral differences here explain the reduced classification accuracy for the 4 cell line validation, where the largest populations of misidentified cells were from the HaCaT-SiHa pairing. A lack of significant spectral differences may suggest that mechanisms behind the biological similarities between these 2 cell lines (eg, their ability to proliferate indefinitely) make a greater contribution to the Raman spectra than those characteristics which constitute a difference (such as DNA content). Significance bands for HaCaT-CaSki and SiHa-CaSki were only seen for the modulated Raman dataset, again suggesting that the application of PCA to differential spectra provides more

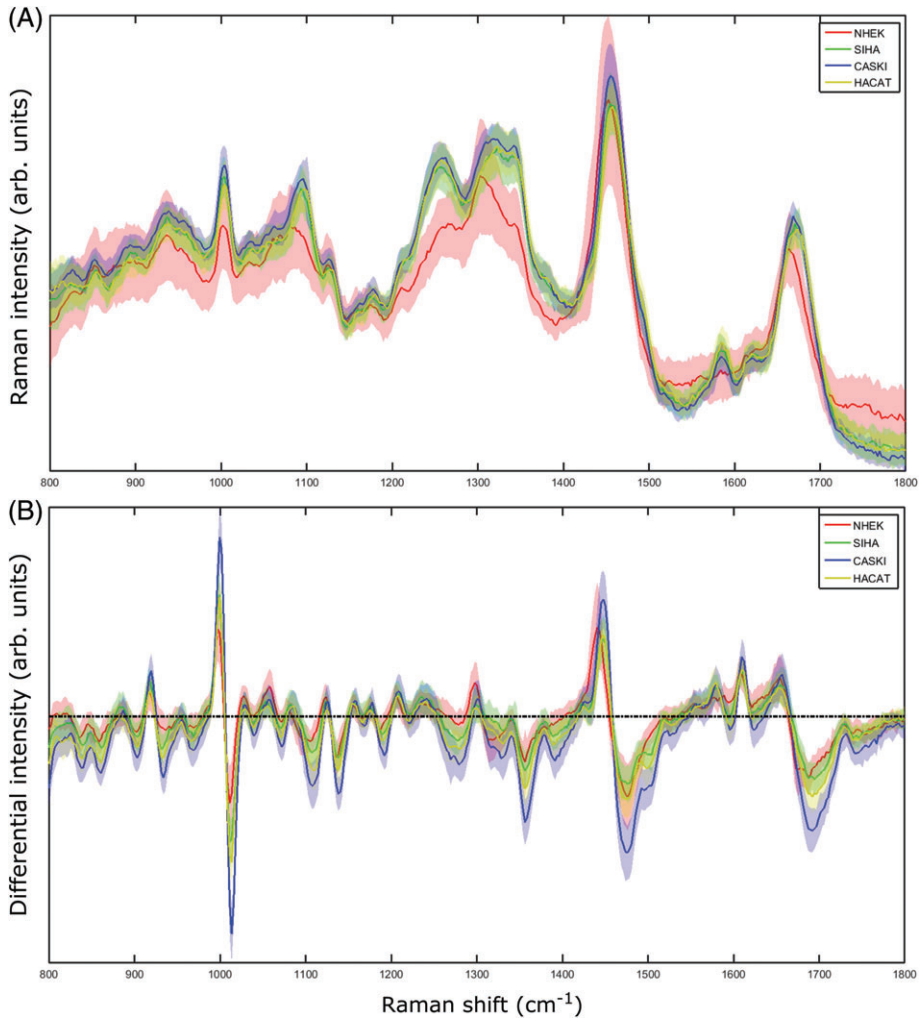


FIGURE 4 (A) Standard and (B) modulated Raman spectral averages for the 4 cell lines used, 800 to 1800 cm^{-1}

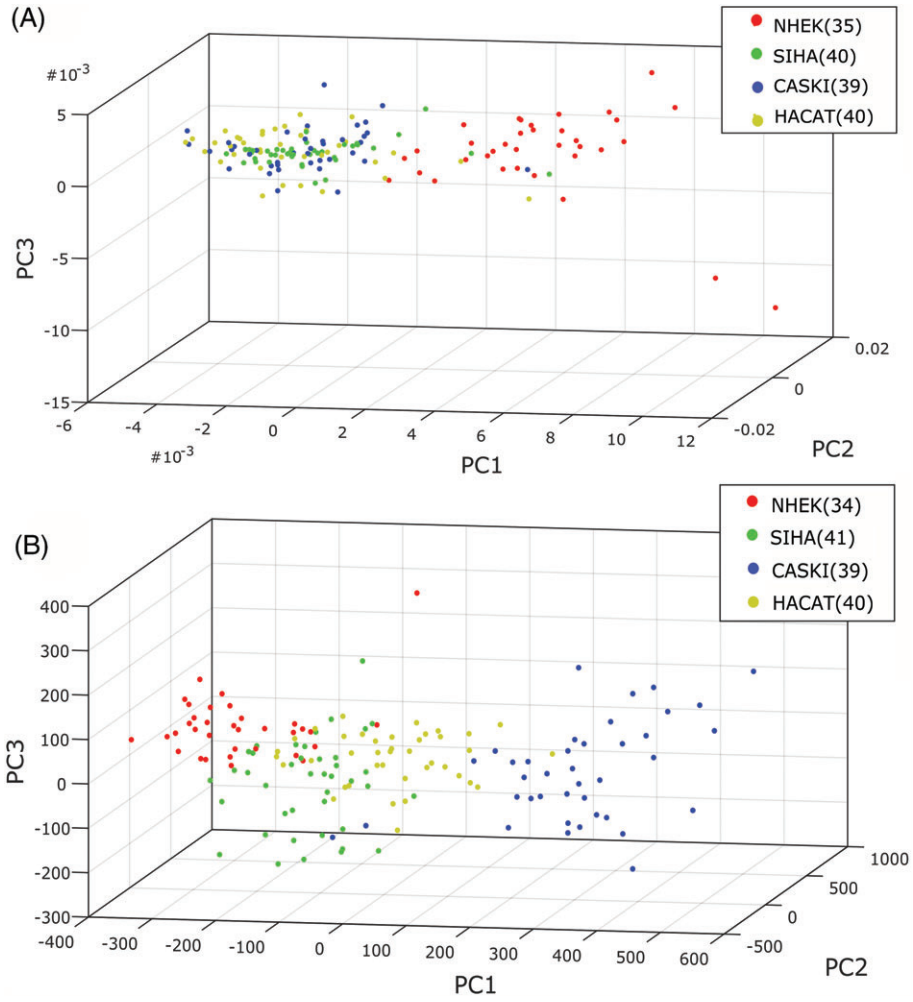


FIGURE 5 (A) Standard and (B) modulated PCA weightings for the 4 cell lines used. Variation in the number of datapoints used arises from the rejection of different spectra in each set

useful information than traditional approaches. Prior to consideration of significance bands, it can be seen in the standard Raman spectra that for each comparison of NHEKs to the 3 established cell lines, the same peaks are demonstrating significant change. The peak at approximately 1590 cm^{-1} (adenine/guanine) only appears for the established cell lines, although significant difference is seen only for NHEK-HaCaT here. It is interesting that the largest peak at approximately 1450 cm^{-1} (CH2/CH3 deformation of lipids/collagen/other hydrocarbon modes) is significant for NHEK-HaCaT (normal vs immortalised) but not for the standard normal vs. transformed cell types (NHEK against SiHa or CaSki). The modulated NHEK-CaSki comparison shows several significant spectral differences not seen in the remaining comparisons. The statistically significant bands and tentative peak assignments based on previous literature [9, 24] are shown in Appendix S2, Table 2.

Overall, it is clear that the use of a differential spectrum in modulated Raman spectroscopy highlights different statistically significant regions of the spectrum in comparison to standard Raman spectra. Due to the format of the differential spectrum, real Raman peaks correspond to differential zero-crossing points. Therefore differential spectral variation represents subtle alterations in the shape, slope and position of peak 'shoulders' as opposed to absolute peak height and position. Small variations in these parameters result from

TABLE 1 The confusion matrix resulting from cross-validation of standard Raman PCA, with accompanying percentage classification accuracy

	Predicted				
Actual	**NHEK**	**SiHa**	**CaSki**	**HaCaT**	**Class acc. (%)**
NHEK	31	2	1	1	88.6
SiHa	2	25	6	7	62.5
CaSki	1	9	19	10	48.7
HaCaT	3	9	6	22	55.0

TABLE 2 The confusion matrix resulting from cross-validation of modulated Raman PCA, with accompanying percentage classification accuracy

	Predicted				
Actual	**NHEK**	**SiHa**	**CaSki**	**HaCaT**	**Class acc. (%)**
NHEK	29	4	0	1	85.3
SiHa	1	32	0	8	78.0
CaSki	0	1	38	0	97.4
HaCaT	0	2	0	38	95.0

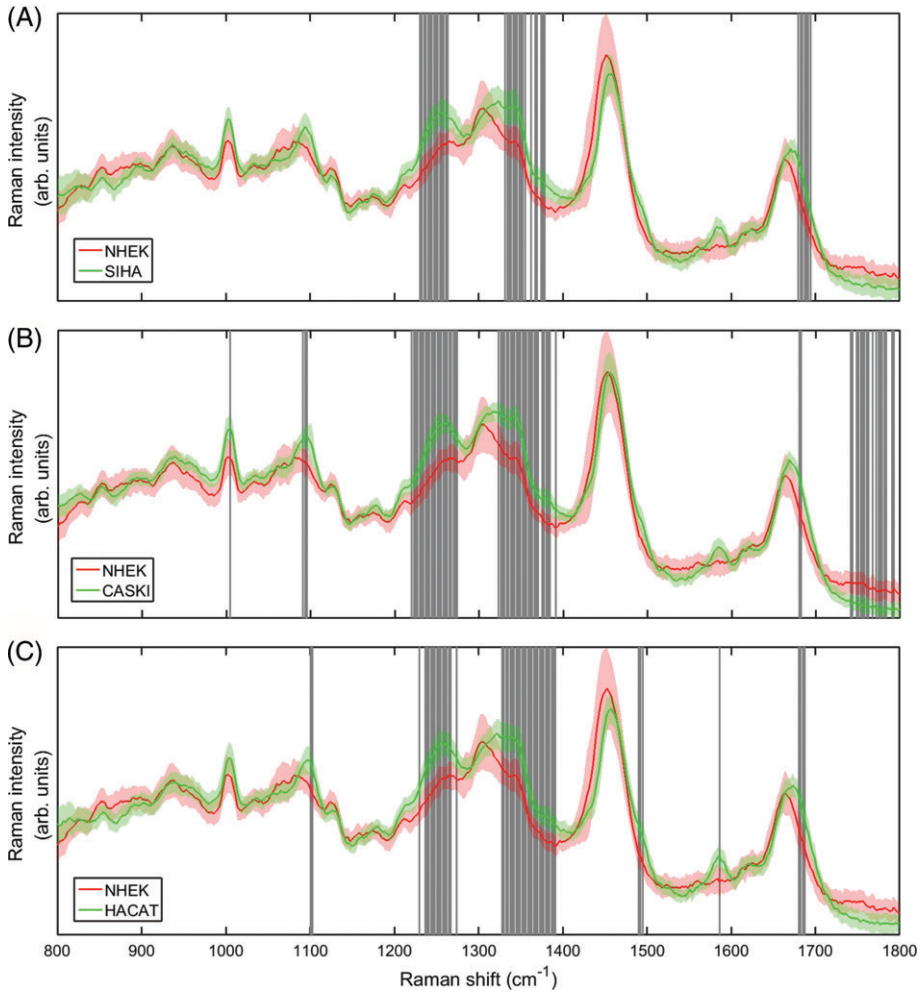


FIGURE 6 Direct comparison of the standard spectra for NHEK against the remaining established cell lines SiHa, CaSki and HaCaT. Shaded bands refer to regions of statisitically significant difference (student's . test) between the spectral intensities

changes to the population of a molecular group (in the case where multiple Raman peaks occupy the same region) or changes to the environment of a bond. This may lead to a significant increase in utility for the differential spectrum approach in biological contexts, as Raman peaks frequently overlap and biological changes are often the result of alterations in molecular conformation in addition to type (as is the case for protein conformation or DNA packaging). The only regions which were statistically significant for all standard Raman spectral comparisons were those corresponding to Amide I:β-sheet/C-O stretch/ν(C-C) and Amide III:α-helix/collagen/CH in phospholipids/C-C in fatty acids. It is suggested that the increased density of protein conformational changes may arise from histones [9]. Many of the remaining significant peak differences also result from DNA-associated structures, including DNA O-P-O backbone stretching and polynucleotide chains. The only differential region which demonstrated statistically significant variations in each case was that of the symmetric ring breathing mode of phenylalanine and tryptophan.

3.3 | Impact of intracellular variation

It was previously found that the Raman spectra for intracellular locations in CHO-K1 cells were very distinct, with the nucleus providing the optimum signal-to-noise ratio [12]. However, that data was generated from adherent cells grown directly onto glass slides, giving the characteristic flattened morphology associated with cells in culture. This distinctive morphology did not translate to the spectral acquisitions from fixed cells considered in this article: the cell rounding effects associated with trypsinisation and fixing made the consistent identification of intracellular organelles challenging. Typical brightfield images are shown in Appendix S3, Figure 1. To investigate the importance of intracellular location on characterisation, further samples of the Section 3.2 cell lines were fixed and each cell interrogated multiple times at locations designated as cell wall, cytoplasm, nucleus and nucleolus. Typical images are shown in Appendix S3, Figure 2. Cells were rejected if the specified components could not be discerned clearly. The nucleolus was identified as a high-density sphere within the nuclear volume. All 4 datasets were recorded in the modulated setup only and analysed as discussed previously. Figure 8 shows the overall PCA plot from the cells, disregarding laser spot location. The resulting confusion matrix and classifications are detailed in Appendix S3, Table 3. Despite some misclassifications (particularly for NHEKs), the average predictive accuracy is high at 91.2%. The

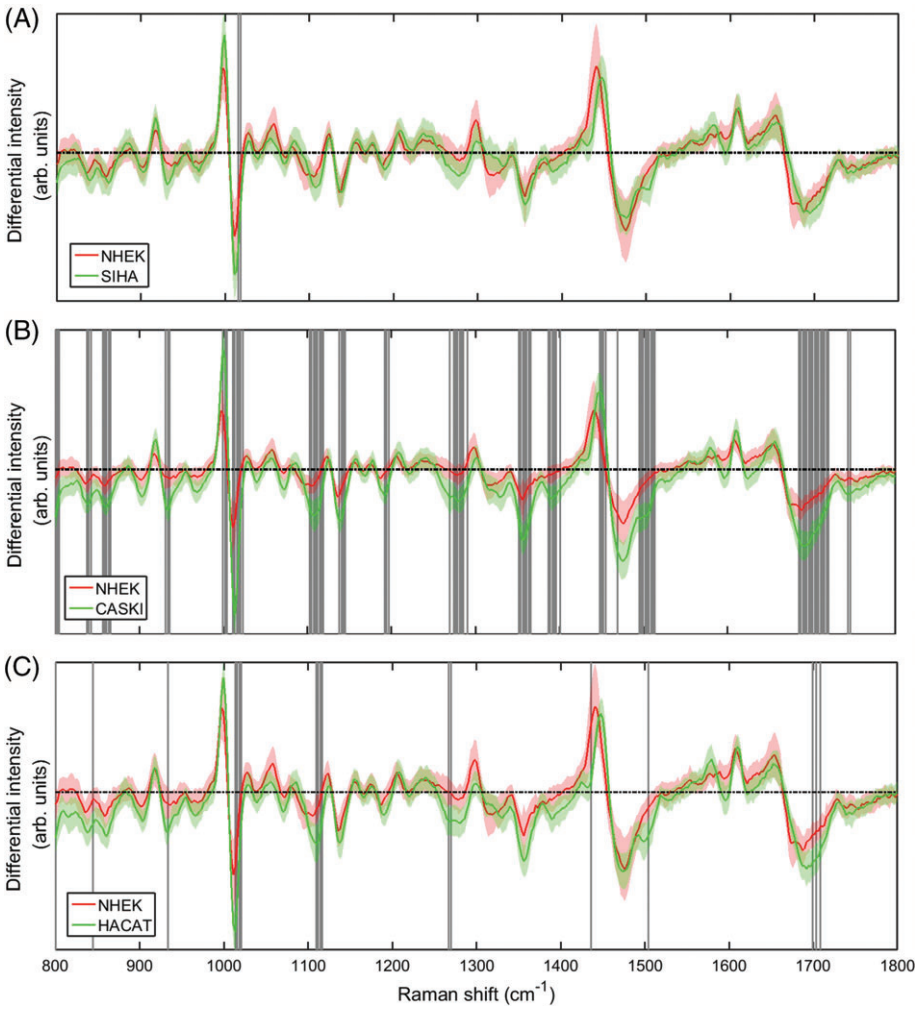


FIGURE 7 Direct comparison of the modulated spectra for NHEK against the remaining established cell lines SiHa, CaSki and HaCaT. Shaded bands refer to regions of statisitically significant difference (student's *t* test) between the spectral intensities

analysis can then be expanded to consider location for each cell line in turn. This can be seen for the SiHa line in Figure 9—the remaining expanded PCA plots can be seen in Appendix S3, Figures 3–5. The classification accuracies for the location-specific subset of each expanded matrix are summarised in Table 3. These are much lower and more variable than both classifications by cell line and location-based results from de Luca et al. experiment [12], suggesting that there is insufficient spectral disparity to separate clusters by location when rounded cells are interrogated in solution.

As such, the matrices were then considered by grouping points according to their location inside (nucleus, nucleolus) or outside (cytoplasm, cell wall) the nucleus. The

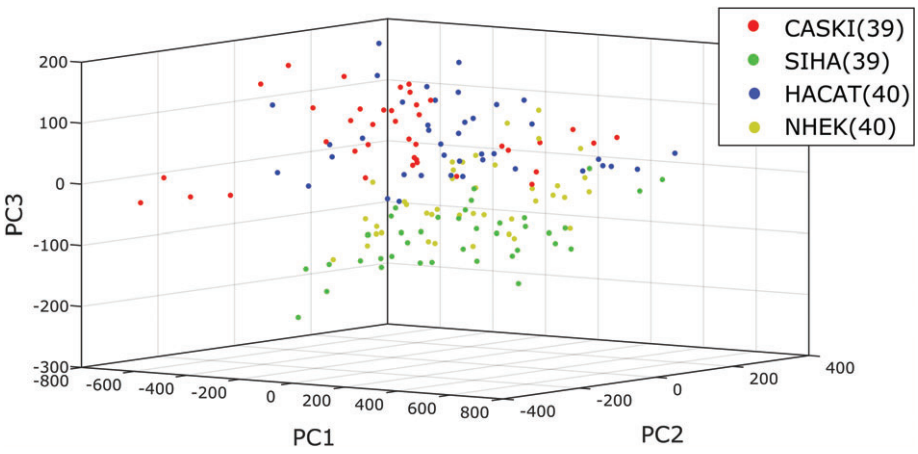


FIGURE 8 The PCA plot for the overall classification of cell lines without consideration of location

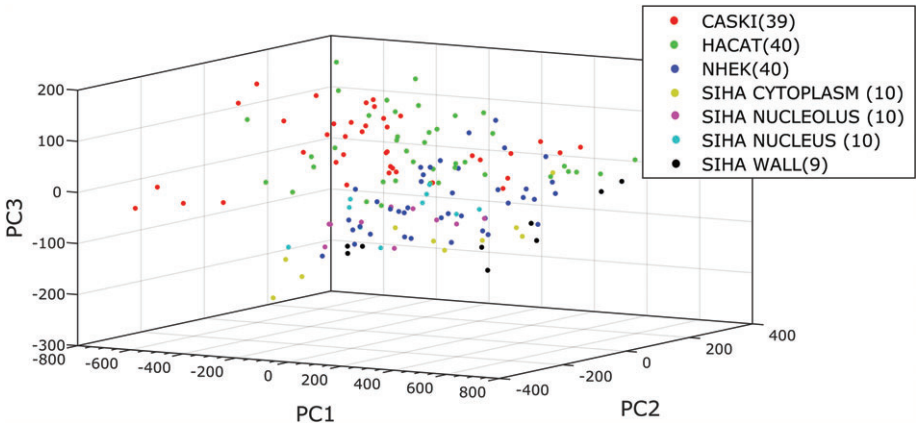


FIGURE 9 The PCA plot for classification where SiHa cells have been considered in terms of sampling by nucleus, nucleolus, cytoplasm or cell wall. The remaining location-expanded plots can be found in Appendix S3

TABLE 3 The percentage classification accuracies for the cell lines by intracellular location. The average classification error is a single SD from the mean

	Classification accuracy (%)				
Loc.	**NHEK**	**SiHa**	**CaSki**	**HaCaT**	**Average**
Nucleus	50	60	50	60	55 ± 5
Nucleolus	50	20	50	80	50 ± 21
Cytoplasm	44	88	44	80	64 ± 20
Cell wall	60	50	60	80	63 ± 11

TABLE 4 The percentage classification accuracies for the cell lines based on location in reference to the nucleus. The average classification error is a single SD from the mean

	Classification accuracy (%)				
Loc.	**NHEK**	**SiHa**	**CaSki**	**HaCaT**	**Average**
Inside	93.3	90.0	90.0	94.1	91.9 ± 1.9
Outside	88.9	86.7	77.8	100.0	88.4 ± 7.9

summarised classification accuracies shown in Table 4. The resulting accuracies from this test indicate that there is sufficient spectral difference to separate these location clusters from inside and outside the nucleus with an accuracy approaching that of cell line classification. However, the fixed cells do not provide sufficient spectral difference to account for specific cellular regions with the accuracy seen for adherent cells interrogated directly on their culture substrates. Thus, cell shape and sampling method should be a key consideration for sampling location accuracy.

3.4 | Impact of other variation sources

Finally, the potential effects of storage time on spectral classification were investigated. Recent work by Hobro et al. considered spectral differences between live mouse embryonic fibroblast cells and those fixed with different agents, including methanol [25]. Methanol acts as both a fixing and permeabilisatio n agent which dehydrates cells and precipitates proteins. In that case, a significant loss of overall Raman signal was observed in cells fixed with methanol, which were mostly associated with lipid bands. The effects of storage in fixative were not considered by Hobro et al. Based on the routine morphological and genomic analysis of stored smear samples without alterations to data, we hypothesised that changes over time after initial fixation would not be significant. If such changes do occur, it would have considerable impact on the use of Raman spectroscopy as a tool for clinical sample classification where samples may be archived in fixing solution for long periods of time either before analysis, for revisiting borderline samples or for comparing with new samples for disease progression. A total of 150 cells were interrogated from SiHa cell samples which had been stored at −20°C for 3 days, 1 month or 6 months in PreservCyt before following the same protocols described previously. The average Raman spectra are in Figure 10. Pairwise comparisons of each sample type as

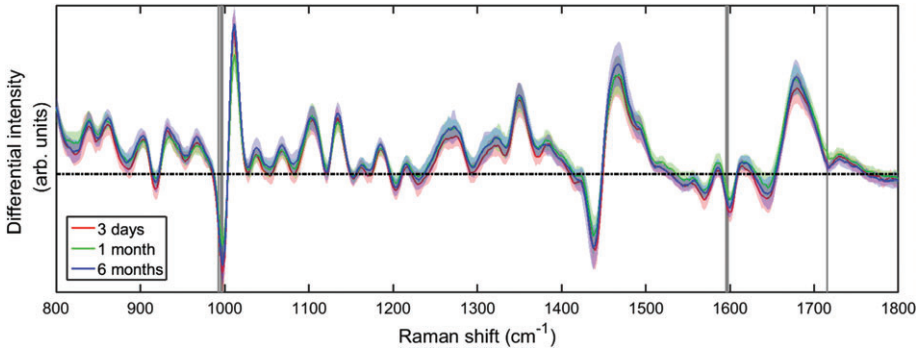


FIGURE 10 Average spectra of 50 cells from each sample, stored in PreservCyt for 3 days, 1 month or 6 months

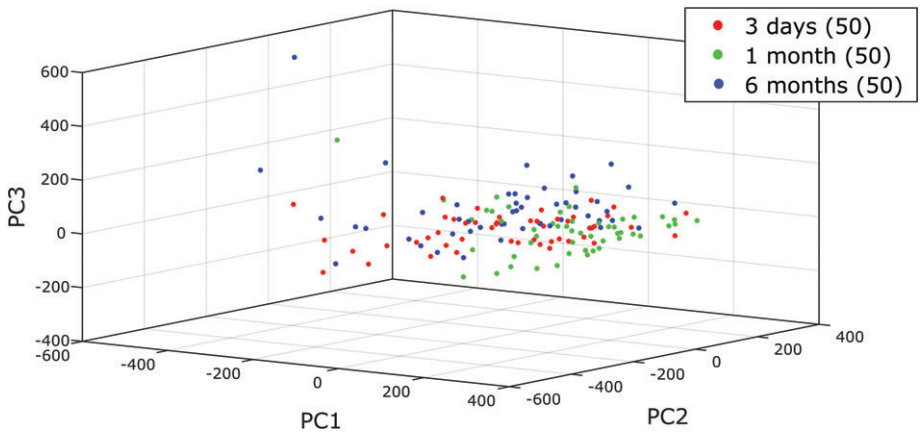


FIGURE 11 The PCA plot of the 3 'archived' fixed samples, stored for different periods of time

TABLE 5 The confusion matrix resulting from cross-validation of the fixed cell storage time dataset, with accompanying percentage classification accuracy

	Predicted			
Actual	**3 d**	**1 mo**	**6 mo**	**Class acc. (%)**
3 d	29	0	21	58.0
1 mo	1	47	2	94.0
6 mo	17	1	32	64.0

conducted before revealed significant differences between 3 days and 1 month cells, and between 1 and 6 months cells, but not between the 1 day and 6 months cells. These differences occur in both the former cases at 992 to 997 cm^{-1} (ring breathing of tryptophan/phenylalanine), 1595 cm^{-1} (adenine/guanine) and 1715 cm^{-1} (C=O, lipids). None of these bands were indicated in the live-fixed cell comparison for methanol [25]. Figure 11 shows the corresponding PCA plot, where grouping can be seen clearly for the 3 days and 6 months clusters in the PC1/PC2 and PC2/PC3 projections. The confusion matrix (Table 5) for the 3 days and 6 months samples are consistent with the hypothesis that no changes should occur during storage, as classification accuracy is low. However, the 1 month sample shows high-classification accuracy, which indicates that there are sufficient spectral differences to classify these nominally identical cells in relation to the remaining samples. The lack of trend according to storage time suggests that while storage should not be an issue for classification, other sources of variation are present and can be detected by the WMRS.

4 | DISCUSSION

In Sections 3.1 and 3.2, the comparisons of standard and modulated methods for spectral acquisition clearly demonstrate that WMRS is able to provide superior cell classification accuracy. The ability to discriminate between HPV positive and negative cervical carcinoma cell lines is an encouraging step towards the potential integration of HPV and smear testing using optical discrimination. However, it must be noted that these carcinoma lines represent different mechanisms of origin, and by extension molecular pathologies. The ability to also identify HPV status in normal cervical cells would be a requirement of an integrated platform and so this will require consideration. The presence of significant spectral differences only in the modulated Raman acquisition between SiHa and CaSki cells, which represent primary and metastatic cervical carcinoma respectively, also points towards greater clinical utility for the sampling method and resulting spectral analysis. The investigation of location impact for spectral sampling provides important data regarding the expected differences between sampling particular cell morphologies or methods of collection. It is suggested that the lower specificity of sampling location for fixed cells in suspension compared to previous experiments with adherent, plated cells is due to the superposition of spectral data from different components of the cell, as the surface area occupied by the cell is lower. This lower margin of error could provide benefits for coupling the spectrometer with automated cell sampling capabilities, and will have to be taken into consideration for clinical samples acquired with a cytology brush. Finally, although overall cell classification was unaffected by storage time in fixative, the relationship between spectral characteristics and fixation was complex, consistent with recent evidence that methanol fixatives affect spectral data, particularly in relation to lipid peaks [25]. The sensitivity of the technique to these changes suggests that similar issues with sample heterogeneity will have to be addressed with large clinical datasets.

5 | CONCLUSION

In this study, fluorescence-free wavelength modulated Raman spectroscopy has been used to consider the benefits and practicalities of such a system in the context of label-free classification of fixed cell samples, such as those from cervical screening. The use of wavelength modulation in

Raman spectroscopy and the resulting removal of autofluorescence provides clear benefits to its ability to automatically classify cells over conventional Raman scattering, without the requirement for manual baseline fitting, additional lasers or optical gates. It is entirely feasible that the technique could be combined with other imaging methods such as optical coherence tomography [26] or digital holographic microscopy [27] in order to provide a cell 'topography' map for automated point sampling. It is hoped that such a system could be introduced to clinical settings in order to reduce the burden on manual cytological diagnosis, as well as increasing the likelihood of a correct outcome for patients.

ACKNOWLEDGMENT

This work was supported by the Engineering and Physical Sciences Research Council and Medical Research Council (EP/L016559/1, EP/P030017/1), and CRUK (A18075 Core Award). We acknowledge M Squared Lasers Ltd for the use of a Solstis laser. Thanks to Professor Ted Hupp for the provision of technical advice and reagents for primary cell culture. We would also like to thank Ramya Bhatia for fixed cell sample donations from the Scottish HPV Archive, and Jan Irvine for facilitating collaboration and research visits.

ORCID

Mingzhou Chen http://orcid.org/0000-0002-6190-5167
C. Simon Herrington http://orcid.org/0000-0001-9177-8165

REFERENCES

[1] CRUK, Screening, http://www.cancerresearchuk.org/about-cancer/cervical-cancer/getting-diagnosed/screening. (accessed: 14 April 2017).
[2] CRUK, 2012 Review of the UK breast screening programme, http://www.cancerresearchuk.org/about-cancer/breast-cancer/screening/screening-2012-review. (accessed: 14 April 2017).
[3] J. H. Smith, J. Patnick, *Achievable Standards, Benchmarks for Reporting, and Criteria for Evaluating Cervical Cytopathology*, 3rd ed., NHS Cancer Screening Programmes, Sheffield **2013**.
[4] P. D. Palma, P. G. Rossi, G. Collina, A. M. Buccoliero, B. Ghiringhello, M. Lestani, G. Onnis, D. Aldovini, G. Galanti, G. P. Casadei, M. Aldi, V. Gomes, P. Giubilato, G. Ronco, and The NTCC pathology group, *Am. J. Clin. Pathol.* **2008**, *129*, 75.
[5] A. L. Mitchell, K. B. Gajjar, G. Theophilou, F. L. Martin, P. L. Martin-Hirsch, *J. Biophotonics* **2014**, *7*, 153.
[6] H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, P. L. Martin-Hirsch, M. J. Walsh, M. R. McAinsh, N. Stone, F. L. Martin, *Nat. Protoc.* **2016**, *11*, 664.
[7] D. Lin, S. Feng, J. Pan, Y. Chen, J. Lin, G. Chen, S. Xie, H. Zeng, R. Chen, *Opt. Express* **2011**, *19*, 13565.
[8] P. R. T. Jess, M. Mazilu, K. Dholakia, A. C. Riches, C. S. Herrington, *Int. J. Cancer* **2009**, *124*, 376.
[9] P. R. T. Jess, D. D. W. Smith, M. Mazilu, K. Dholakia, A. C. Riches, C. S. Herrington, *Int. J. Cancer* **2007**, *121*, 2723.
[10] I. R. M. Ramos, A. Malkin, F. M. Lyng, *Biomed. Res. Int.* **2015**, *2015*, 561242.
[11] E. Vargis, E. M. Kanter, S. K. Majumder, M. D. Keller, R. B. Beaven, G. G. Rao, A. Mahadevan-Jansen, *Analyst* **2011**, *136*, 2981.
[12] A. C. De Luca, M. Mazilu, A. Riches, C. S. Herrington, K. Dholakia, *Anal. Chem.* **2010**, *82*, 738.
[13] D. Wei, S. Chen, Q. Liu, *Appl. Spectrosc. Rev.* **2015**, *50*, 387.
[14] B. B. Praveen, M. Mazilu, R. F. Marchington, C. S. Herrington, A. Riches, K. Dholakia, *PLoS ONE* **2013**, *8*, e67211.
[15] M. Mazilu, A. C. De Luca, A. Riches, C. S. Herrington, K. Dholakia, *Opt. Express* **2010**, *18*, 11382.
[16] M. Chen, N. McReynolds, E. C. Campbell, M. Mazilu, J. Barbosa, K. Dholakia, S. J. Powis, *PLoS ONE* **2015**, *10*, e0125158.
[17] E. Canetta, M. Mazilu, A. C. De Luca, A. E. Carruthers, K. Dholakia, S. Neilson, H. Sargeant, T. Briscoe, C. S. Herrington, A. C. Riches, *J. Biomed. Opt.* **2011**, *16*, 037002.
[18] H. Abdi, L. J. Williams, *WIREs Comp. Stat.* **2010**, *2*, 433.
[19] B. M. Wise, N. L. Ricker, D. F. Veltkamp, B. R. Kowalski, *Process Control and Quality* **1990**, *1*, 41.
[20] R. Bhatia, The Scottish HPV Archive (report), http://www.ed.ac.uk/pathology/research/scottish-hpv-archive. (accessed: 14 April 2017).
[21] P. Boukamp, R. T. Petrussevska, D. Breitkreutz, J. Hornung, A. Markham, N. E. Fusenig, *J. Cell Biol.* **1988**, *106*, 761.
[22] F. Friedl, I. Kimura, T. Osato, Y. Ito, *Proc. Soc. Exp. Biol. Med.* **1970**, *135*, 543.
[23] R. A. Pattillo, R. O. Hussa, M. T. Story, A. C. Ruckert, M. R. Shalaby, R. F. Mattingly, *Science* **1977**, *196*, 1456.
[24] A. C. S. Talari, Z. Movasaghi, S. Rehman, I. ur Rehman, *Appl. Spectrosc. Rev.* **2015**, *50*, 46.
[25] A. J. Hobro, N. I. Smith, *Vib Spectrosc* **2017**, *91*, 31.
[26] M. Chen, J. Mas, L. H. Forbes, M. R. Andrews, K. Dholakia, *J. Biophotonics*, **2017**, e201700129.
[27] N. McReynolds, F. G. M. Cooke, M. Chen, S. J. Powis, K. Dholakia, *Sci. Rep.* **2017**, *7*, 43631.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Appendix S1. Details of the origin and cell culture reagents.

Appendix S2. Statistically significant Raman spectral regions.

Appendix S3. Confusion matrix and classification accuracy

How to cite this article: Woolford L, Chen M, Dholakia K, Herrington CS. Towards automated cancer screening: Label-free classification of fixed cell samples using wavelength modulated Raman spectroscopy. *J. Biophotonics*. 2018;11:e201700244. https://doi.org/10.1002/jbio.201700244

Chapter 9

Appendix B: SERS-SRS imaging

To date, published examples of SRS imaging with additional nanoparticle-based labelling are limited. Earlier work indicated that the short laser pulses - and therefore high peak powers - typically employed with SRS in order to excite third-order processes were insufficient for sustaining LSPRs on colloidal metal due to the intensity fluctuations associated with the phenomenon [264] and also cause significant photodegradation [265]. Continuous-wave SRS has shown greater promise [266, 267] and some studies have been performed on supporting LSPRs with short pulses [268] but the field is still in its infancy.

9.1 Materials and methods

9.1.1 Stimulated SERS mapping

Stimulated Raman spectroscopy (SRS) images were acquired using a custom-built multi-modal microscope setup, designed by Dr Alan Serrels as part of a collaboration with the Frame, Brunton and Elfick groups at the University of Edinburgh (Figure 9.1). A picoEmerald (APE) laser system provided both a tunable pump laser (720-990nm, 7ps pulse) and a spatially overlapped Stokes laser (1064nm, 5–6ps), set to 80MHz repetition rate. The laser was inserted into a multi-photon laser-scanning microscope (FV1000MPE) coupled to an objective lens (NA 1.05, XLPL25XWMP) using a short-pass 690nm dichroic mirror. For SRS measurements, the Stokes beam was modulated with a 20MHz electro-optic modulator (EoM). Forward scattered light was collected by a 25$\times$ water-immersion objective lens (NA 1.00, XLUMPLFLN) and the Stokes scattering removed using an ET890/220m filter (Chroma). A telescope focused the light onto an APE silicon photodiode connected to an APE lock-in amplifier which was fed into the analogue unit of the microscope.

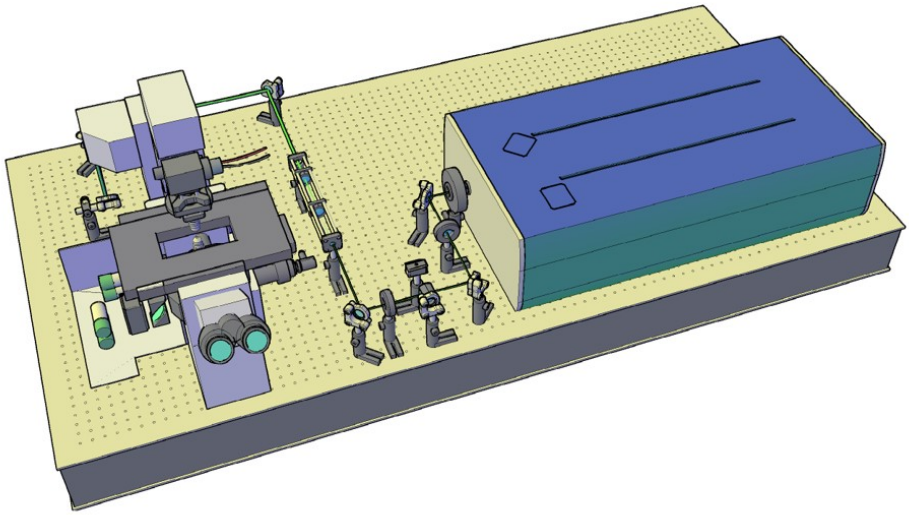


FIGURE 9.1: The optical setup for the multi-photon laser scanning microscope (left) coupled to picoEmerald laser system (blue and green box, right).

The pump laser was tuned to 810.5nm for the 2939.6cm^{-1} protein-rich CH_3 stretch to delineate cell membranes, before being tuned between 905.0 and 911.1nm, equivalent to 1577.2-1651.2cm^{-1} with the Stokes laser used. The equation for pump wavelength to wavenumber conversion is given below, where the wavenumber $\nu = \frac{1}{\lambda} \times 10^7$. The laser powers after the objective were measured to 40-70mW for the pump laser and up to 70mW for the Stokes laser before imaging. All images were recorded at 800×800 pixels, equivalent to a 127.041µm square, with a pixel dwell time of 20µs. Olympus FV10-ASW software was used for image acquisition and image analysis and processing was performed in ImageJ.

$$Pump_\nu - Stokes_\nu = Interrogated_\nu \tag{9.1}$$

9.2 Consideration of stimulated SERS imaging

In order to address the continuing issue of cumulative acquisition time for the practical imaging of cervical smear samples, another imaging modality was considered:

stimulated Raman spectroscopy (SRS). As discussed in the introduction to this Chapter, SRS is able to produce maps of intrinsic Raman vibrational modes, creating digital analogues of tissue and cell stains. It was hypothesised that by combining SERS-active nanoparticle labelling with global mapping, a 'hybrid' imaging system could be created which would provide the targeted analysis of a p16-specific conjugate alongside global structural information about the cells themselves. This could be used for a data-rich exploration of structural biology in the cervical cancer context. More pragmatically, spatial information about the cell boundaries would provide a straightforward method of determining the quantity of nanoparticles inside and outside the cells which could not be achieved easily using the relatively low-contrast brightfield images seen in Section 5.5.2. Furthermore, SRS imaging simply provides faster image acquisition, with pixel dwell times of 20µs as opposed to 1s.

The fixed cell samples prepared for SERS imaging in Section 5.5.2 were imaged using the SRS system as described in Section 9.1.1 and published previously by Serrels et al. [269]. The results can be seen in Figure 9.2. Although the images show signal in the region of the 1606 and 1636cm^{-1} BPE peaks, this signal does not vary with peak intensity as the pump laser sweeps through the wavenumber region. Thus, the signal can only be attributed to background or cellular vibrational scattering. Given the stability requirements outlined in Lee et al's paper [267], it is unsurprising that the in-house colloids were unable to support LSPRs with such a short laser pulse. A continuous-wave optical setup was not available to test in the SERS-SRS context and so it was suggested that alkyne labelling of antibodies should be explored instead, as described in Section 5.6.

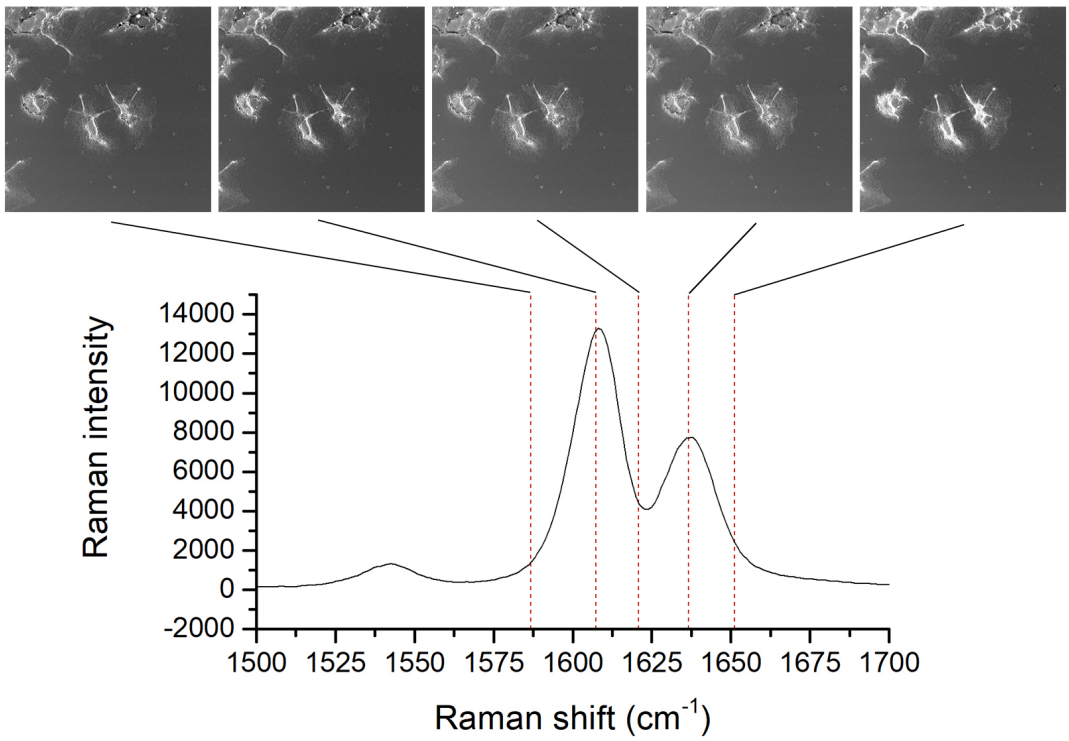


FIGURE 9.2: Representative images of Raman shift sweep of targeted conjugates in fixed SiHa w/t cells as prepared for standard SERS experiments. The image intensity does change with shift but not in proportion to the Raman signal of BPE as shown, meaning that the SRS setup is not capable of supporting hyrbid SERS imaging.

Bibliography

[1] Elizabeth Vargis et al. "Near-infrared Raman Microspectroscopy Detects High-risk Human Papillomaviruses". In: *Translational Oncology* 5.3 (2012), pp. 172–179. ISSN: 1936-5233. DOI: http://dx.doi.org/10.1593/tlo.12106. URL: http://www.sciencedirect.com/science/article/pii/S1936523312800052.

[2] *Blueprints Obstetrics & Gynecology*. Lippincott Williams & Wilkins, 2013.

[3] Diama Bhadra Vale, Joana Froes Bragança, and Luiz Carlos Zeferino. "Cervical Cancer Screening in Low- and Middle-Income Countries". In: *Uterine Cervical Cancer: Clinical and Therapeutic Perspectives*. Ed. by Samir A. Farghaly. Cham: Springer International Publishing, 2019, pp. 53–59. ISBN: 978-3-030-02701-8. DOI: 10.1007/978-3-030-02701-8_4.

[4] Rosliza Abdul Manaf, Suriani Ismail, and Nwabichie Chinemerem Cecilia. "Global Burden of Cervical Cancer: A Literature Review". In: *International Journal of Public Health and Clinical Sciences* April (2017).

[5] Global Burden of Disease Cancer Collaboration. "Europe PMC Funders Group The Global Burden of Cancer 2013". In: *JAMA Oncol.* 1.January 2014 (2015), pp. 505–527. DOI: 10.1001/jamaoncol.2015.0735.The.

[6] Jacques Ferlay et al. "Cancer incidence and mortality worldwide : Sources , methods and major patterns in GLOBOCAN 2012". In: *International Journal of Cancer* 386 (2015). DOI: 10.1002/ijc.29210.

[7] Globocan. *Estimated Cancer Incidence, Mortality and Prevalence Worldwide*. 2012. URL: http://globocan.iarc.fr/Pages/Map.aspx\#.

[8] Emmanuela Gakidou, Stella Nordhagen, and Ziad Obermeyer. "Coverage of cervical cancer screening in 57 countries: Low average levels and large inequalities". In: *PLoS Medicine* 5.6 (2008), pp. 0863–0868. ISSN: 15491277. DOI: 10.1371/journal.pmed.0050132.

[9] Ophira Ginsburg et al. "The global burden of women's cancers: a grand challenge in global health". In: *The Lancet* 389.10071 (2017), pp. 847–860. ISSN: 1474547X. DOI: 10.1016/S0140-6736(16)31392-7. URL: http://dx.doi.org/10.1016/S0140-6736(16)31392-7.

[10] Murat Gultekin et al. "Initial results of population based cervical cancer screening program using HPV testing in one million Turkish women". In: *International Journal of Cancer* 1958 (2018), pp. 1952–1958. DOI: 10.1002/ijc.31212.

[11] Sheri A Denslow et al. "Incidence and progression of cervical lesions in women with HIV: A systematic global review". In: *Int J STD AIDS* 25.3 (2014), pp. 163–177. DOI: 10.1177/0956462413491735.Incidence.

[12] *Scottish Cervical Screening Programme Statistics 2016-17*. Tech. rep. September. NHS Scotland, 2017.

[13] J T Schiller, P M Day, and R C Kines. "Current understanding of the mechanism of HPV infection". In: *Gynecologic Oncology* 118.1 Suppl (2010), S12–7. DOI: 10.1016/j.ygyno.2010.04.004.Current.

[14] *Cancer Chemotherapy: Basic Science to the Clinic*. Wiley, 2008.

[15] E Burd. "Human papillomavirus and cervical cancer". In: *Clin Microbiol Rev* 16.1 (2003), pp. 1–17. ISSN: 0893-8512. DOI: 10.1128/CMR.16.1.1. URL: http://www.sciencedirect.com/science/article/pii/S0140673607614160.

[16] John Doorbar. "Molecular biology of human papillomavirus infection and cervical cancer." In: *Clinical science (London, England : 1979)* 110.5 (2006), pp. 525–41. ISSN: 0143-5221. DOI: 10.1042/CS20050369. URL: http://www.clinsci.org/content/110/5/525.abstract.

[17] Inês Raquel et al. "Current Advances in the Application of Raman Spectroscopy for Molecular Diagnosis of Cervical Cancer". In: *BioMed Research International* 2015 (2015).

[18] Corina Ziegert et al. "A comprehensive analysis of HPV integration loci in anogenital lesions combining transcript and genome-based amplification techniques." In: *Oncogene* 22.25 (2003), pp. 3977–3984. ISSN: 09509232. DOI: 10.1038/sj.onc.1206629.

[19] Teresa Kim, Samer N Khader, and D Yitzchak Goldstein. "Educational Case : Cervical Neoplasia : HPV and Its Link to Cancer". In: *Academic Pathology* 5 (2018), pp. 1–6. DOI: 10.1177/2374289518770651.

[20] Ross L Cameron et al. "Human Papillomavirus Prevalence and Herd Immunity after Introduction of Vaccination". In: *Emerging Infectious Diseases* 22.1 (2016), pp. 2009–2013.

[21] David Mesher et al. "Continuing reductions in HPV 16 / 18 in a population with high coverage of bivalent HPV vaccination in England : an ongoing cross-sectional study". In: (2016), pp. 1–8. DOI: 10.1136/bmjopen-2015-009915.

[22] Tim Palmer et al. "Prevalence of cervical disease at age 20 after immunisation with bivalent HPV vaccine at age 12-13 in Scotland : retrospective population study". In: *British Medical Journal* June (2019), pp. 1–10. DOI: 10.1136/bmj.l1161.

[23] T J Palmer et al. "HPV immunisation and cervical screening — confirmation of changed performance of cytology as a screening test in immunised women : a retrospective population-based cohort study". In: *British journal of cancer* 114.5 (2016), pp. 582–589. ISSN: 0007-0920. DOI: 10.1038/bjc.2015.474. URL: http://dx.doi.org/10.1038/bjc.2015.474.

[24] Ping Teresa Yeh et al. "Self-sampling for human papillomavirus (HPV) testing : a systematic review and meta-analysis". In: *BMJ Global Health* (2019). DOI: 10.1136/bmjgh-2018-001351.

[25] CRUK. *Cervical cancer statistics*. 2019. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/cervical-cancer#heading-Zero.

[26] Mert Ulas Barut and Ahmet Kale. "Analysis of Sensitivity , Specificity , and Positive and Negative Predictive Values of Smear and Colposcopy in Diagnosis of Premalignant and Malignant Cervical Lesions". In: *Medical Science Monitor* (2015), pp. 3860–3867. DOI: 10.12659/MSM.895227.

[27] M G Dijkstra et al. "Cervical cancer screening : on the way to a shift from cytology to full molecular screening". In: (2014), pp. 381–385. DOI: 10.1093/annonc/mdt538.

[28] Paolo Dalla Palma et al. "The risk of false-positive histology according to the reason for colposcopy referral in cervical cancer screening: A blind revision of all histologic lesions found in the NTCC trial". In: *American Journal of Clinical Pathology* 129.1 (2008), pp. 75–80. ISSN: 00029173. DOI: 10.1309/EWYGWFRRM8798U5P.

[29] Saskia Bulk et al. "Cervical cancer in the Netherlands 1989 – 1998 : Decrease of squamous cell carcinoma in older women , increase of adenocarcinoma in younger women". In: *International Journal of Cancer* 1009.August 2004 (2005), pp. 1005–1009. DOI: 10.1002/ijc.20678.

[30] Ahti Anttila, Eero Pukkala, and Matti Hakama. "EFFECT OF ORGANISED SCREENING ON CERVICAL CANCER INCIDENCE AND MORTALITY IN FINLAND , 1963 – 1995 : RECENT INCREASE IN CERVICAL". In: *International Journal of Cancer* 65.May (1999), pp. 59–65.

[31] Matejka Rebolj et al. "Primary cervical screening with high risk human papillomavirus testing : observational study". In: *British Medical Journal* (2019). DOI: 10.1136/bmj.l240.

[32] Anthony E Rizzo. "Update on primary HPV screening for cervical cancer prevention". In: *Current Problems in Cancer* 42.5 (2018), pp. 507–520. ISSN: 0147-0272. DOI: 10.1016/J.CURRPROBLCANCER.2018.06.013. URL: https://www.sciencedirect.com/science/article/abs/pii/S0147027218301727?via{\%}3Dihub.

[33] Guglielmo Ronco et al. "Efficacy of HPV-based screening for prevention of invasive cervical cancer: Follow-up of four European randomised controlled trials". In: *The Lancet* 383.9916 (2014), pp. 524–532. ISSN: 1474547X. DOI: 10.1016/S0140-6736(13)62218-7.

[34] Dorien C. Rijkaart et al. "Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: Final results of the POBASCAM randomised controlled trial". In: *The Lancet Oncology* 13.1 (2012), pp. 78–88. ISSN: 14702045. DOI: 10.1016/S1470-2045(11)70296-0.

[35] Guglielmo Ronco et al. "Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial". In: *The Lancet Oncology* 11.3 (2010), pp. 249–257. ISSN: 14702045. DOI: 10.1016/S1470-2045(09)70360-2.

[36] Maaike G. Dijkstra et al. "Safety of extending screening intervals beyond five years in cervical screening programmes with testing for high risk human papillomavirus: 14 year follow-up of population based randomised cohort in the Netherlands". In: *BMJ (Clinical research ed.)* 355 (2016), p. i4924. ISSN: 17561833. DOI: 10.1136/bmj.i4924.

[37] M. Zorzi et al. "A 3-year interval is too short for re-screening women testing negative for human papillomavirus: a population-based cohort study". In: *BJOG: An International Journal of Obstetrics and Gynaecology* 124.10 (2017), pp. 1585–1593. ISSN: 14710528. DOI: 10.1111/1471-0528.14575.

[38] Kate Cuschieri et al. "Eurogin roadmap 2017: Triage strategies for the management of HPV-positive women in cervical screening programs". In: *International Journal of Cancer* 00 (2018).

[39] Dorothy A Machalek et al. "Routine cervical screening by primary HPV testing: early findings in the renewed National Cervical Screening Program". In: *The Medical Journal of Australia* (2019), pp. 113–119. DOI: 10.5694/mja2.50223.

[40] Alana L. Mitchell et al. "Vibrational spectroscopy of biofluids for disease screening or diagnosis: Translation from the laboratory to a clinical setting". In: *Journal of Biophotonics* 7.3-4 (2014), pp. 153–165. ISSN: 18640648. DOI: 10.1002/jbio.201400018.

[41] Ramya Bhatia et al. "Host chemokine signature as a biomarker for the detection of pre-cancerous cervical lesions". In: *Oncotarget* 9.26 (2018), pp. 18548–18558. ISSN: 1949-2553. DOI: https://doi.org/10.18632/oncotarget.24946. URL: https://www.oncotarget.com/article/24946/.

[42] C S Herrington. "What we could do now : molecular pathology of gynaecological cancer". In: *Journal of clinical pathology* March (2001), pp. 222–224.

[43] N Murphy et al. "p16INK4A, CDC6, and MCM5: predictive biomarkers in cervical preinvasive neoplasia and cervical cancer." In: *Journal of clinical pathology* 58.5 (2005), pp. 525–34. ISSN: 0021-9746. DOI: 10.1136/jcp.2004.018895. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1770660{\&}tool=pmcentrez{\&}rendertype=abstract.

[44] Nicolas Wentzensen and Magnus von Knebel Doeberitz. "Biomarkers in cervical cancer screening". In: *Disease Markers* 23 (2007), pp. 315–330.

[45] John Doorbar. "Papillomavirus life cycle organization and biomarker selection." In: *Disease markers* 23.4 (2007), pp. 297–313. ISSN: 0278-0240. DOI: 10.1155/2007/613150.

[46] Francesca Carozzi et al. "Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: a nested substudy of the NTCC randomised controlled trial". In: *The Lancet Oncology* 9.10 (2008), pp. 937–945. ISSN: 14702045. DOI: 10.1016/S1470-2045(08)70208-0.

[47] Heather Griffin et al. "Stratification of HPV-induced cervical pathology using the virally encoded molecular marker E4 in combination with p16 or MCM". In: *Modern pathology* (2015), pp. 977–993. ISSN: 0893-3952. DOI: 10.1038/modpathol.2015.52.

[48] Heather Griffin et al. "E4 Antibodies Facilitate Detection and Type-Assignment of Active HPV Infection in Cervical Disease". In: *PLoS ONE* 7.12 (2012). ISSN: 19326203. DOI: 10.1371/journal.pone.0049974.

[49] Young Tae Kim and Min Zhao. "Aberrant cell cycle regulation in cervical carcinoma". In: *Yonsei Medical Journal* 46.5 (2005), pp. 597–613. ISSN: 05135796. DOI: 10.3349/ymj.2005.46.5.597.

[50] Ruediger Klaes et al. "OVEREXPRESSION OF p16 INK4A AS A SPECIFIC MARKER FOR DYSPLASTIC AND NEOPLASTIC EPITHELIAL CELLS OF THE CERVIX UTERI". In: *International Journal of Cancer* 284.July 2000 (2001), pp. 276–284.

[51] P. Krishnappa et al. "Expression of P16 in high-risk human papillomavirus related lesions of the uterine cervix in a government hospital, Malaysia". In: *Diagnostic pathology* (2014).

[52] Jiaying Lin et al. "Prognostic significance of overexpressed p16INK4a in patients with cervical cancer: a meta-analysis." In: *PloS one* 9.9 (2014), e106384. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0106384. URL: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0106384.

[53] Michele Follen et al. "Cervical cancer chemoprevention, vaccines, and surrogate endpoint biomarkers." In: *Cancer* 98.9 Suppl (2003), pp. 2044–51. ISSN: 0008-543X. DOI: 10.1002/cncr.11674. URL: http://www.ncbi.nlm.nih.gov/pubmed/14603541.

[54] Maurice Canham et al. "Increased cycling cell numbers and stem cell associated proteins as potential biomarkers for high grade human papillomavirus+ve pre-neoplastic cervical disease". In: *PLoS ONE* 9.12 (2014), pp. 1–18. ISSN: 19326203. DOI: 10.1371/journal.pone.0115379.

[55] *Molecular Pathology: The Molecular Basis of Human Disease*. Elsevier, 2018.

[56] J Louise Jones et al. "Morphomolecular pathology : setting the framework for a new generation of pathologists". In: *British journal of cancer* 117.11 (2017), pp. 1581–1582. ISSN: 0007-0920. DOI: 10.1038/bjc.2017.340. URL: http://dx.doi.org/10.1038/bjc.2017.340.

[57] Ian Cree. *The pathology of cancer: From stained slides to artificial intellgence*. 2016. URL: https://www.rcpath.org/discover-pathology/news/blog-the-pathology-of-cancer-from-stained-slides-to-artificial-intelligence.html.

[58] Manuel Salto-tellez, Jacqueline A James, and Peter W Hamilton. "ScienceDirect Molecular pathology e The value of an integrative approach". In: *Molecular Oncology* 8.0 (2014), pp. 4–9.

[59] Matteo Fassan. "Molecular Diagnostics in Pathology". In: *Arch Pathol Lab Med* 142.March (2018). DOI: 10.5858/arpa.2017-0269-RA.

[60] Min Lin et al. "Recent Advances on the Molecular Mechanism of Cervical Carcinogenesis Based on Systems Biology Technologies". In: *Computational and Structural Biotechnology Journal* 17.109 (2019), pp. 241–250. ISSN: 2001-0370. DOI: 10.1016/j.csbj.2019.02.001. URL: https://doi.org/10.1016/j.csbj.2019.02.001.

[61] Xueqiong Zhu et al. "Proteomic identification of differentially-expressed proteins in squamous cervical cancer". In: *Gynecologic Oncology* 112 (2009), pp. 248–256. DOI: 10.1016/j.ygyno.2008.09.045.

[62] Ming-zhu Yin et al. "Identification of phosphatidylcholine and lysophosphatidylcholine as novel biomarkers for cervical cancers in a prospective cohort study". In: *Tumor Biology* (2016), pp. 5485–5492. DOI: 10.1007/s13277-015-4164-x.

[63] Feng Qiu et al. "New serum biomarker identification and analysis by mass spectrometry in cervical precancerous lesion and acute cervicitis in South China". In: *Cancer Management and Research* (2019).

[64] Yubo Fan et al. "Review Article Screening of Cervical Cancer with Self-Collected Cervical Samples and Next-Generation Sequencing". In: *Disease markers* 2018 (2018). DOI: 10.1155/2018/4826547.

[65] Pornjarim Nilyanimit, Jira Chansaenroj, and Yong Poovorawan. "Comparison of Four Human Papillomavirus Genotyping Methods: Next-generation Sequencing, INNO-LiPA, Electrochemical DNA Chip, and Nested-PCR". In: *Annals of Laboratory Medicine* (2018), pp. 139–146.

[66] Xin Yi et al. "A New PCR-Based Mass Spectrometry System for High- Risk HPV , Part I Methods". In: *Anatomic Pathology* (2011), pp. 913–919. DOI: 10.1309/AJCPWTZDT0Q7DOVI.

[67] Tak Hong Cheung et al. "Liquid biopsy of HPV DNA in cervical cancer". In: *Journal of Clinical Virology* 114.March (2019), pp. 32–36. ISSN: 1386-6532. DOI: 10.1016/j.jcv.2019.03.005. URL: https://doi.org/10.1016/j.jcv.2019.03.005.

[68] D. Hunter Best et al. "Molecular Pathology Methods". In: *Molecular Pathology in Clinical Practice:Second Edition* (2016), pp. 1–1001. DOI: 10.1007/978-3-319-19674-9.

[69] Peter Birner, Gerald Prager, and Berthold Streubel. "Molecular pathology of cancer : how to communicate with disease". In: *ESMO Open Cancer Horizons* (2016), pp. 1–6. DOI: 10.1136/esmoopen-2016-000085.

[70] Dennis J Slamon et al. "Studies of the HER-2/neu Proto-oncogene in Human Breast and Ovarian Cancer". In: *Science* (1989).

[71] Ian A Cree et al. "Guidance for laboratories performing molecular pathology for cancer patients". In: *Journal of clinical pathology* (2014), pp. 923–931. DOI: 10.1136/jclinpath-2014-202404.

[72] Frederick S Nolte. "Hepatitis B and C Viruses". In: *Molecular Pathology in Clinical Practice* (2016), pp. 641–657. DOI: 10.1007/978-3-319-19674-9.

[73] Raymund R Razonable and Randall T Hayden. "Viral Infections in Transplant Recipients". In: *Molecular Pathology in Clinical Practice* (2016). DOI: 10.1007/978-3-319-19674-9.

[74] Kevin Alby and Melissa B Miller. "Bacterial Infections". In: *Molecular Pathology in Clinical Practice* (2016). DOI: 10.1007/978-3-319-19674-9.

[75] Angela M Caliendo and Colleen S Kraft. "Human Immunodeficiency Virus Type 1". In: *Molecular Pathology in Clinical Practice* (2016). DOI: 10.1007/978-3-319-19674-9.

[76] Jeanne A Jordan. "Sexually Transmitted Diseases". In: *Molecular Pathology in Clinical Practice* (2016). DOI: 10.1007/978-3-319-19674-9.

[77] Li Yu et al. "Application of p16 / Ki-67 dual-staining cytology in cervical cancers". In: *Journal of Cancer* 10 (2019). DOI: 10.7150/jca.32743.

[78] Diya Das and Moumita Sengupta. "Role of p16Ki-67 Dual Immunostaining in Detection of Cervical Cancer Precursors". In: *Journal of Cytology* (2018).

[79] Mary Mcmenamin, Michael Mckenna, and Andrew Mcdowell. "Clinical Utility of CINtec PLUS Triage in Equivocal Cervical Cytology and Human Papillomavirus Primary Screening". In: *American Journal of Clinical Pathology* (2018), pp. 512–521. DOI: 10.1093/AJCP/AQY073.

[80] Thomas C Wright et al. "Gynecologic Oncology Triaging HPV-positive women with p16 / Ki-67 dual-stained cytology : Results from a sub-study nested into the ATHENA trial". In: *Gynecologic Oncology* 144.1 (2017), pp. 51–56. ISSN: 0090-8258. DOI: 10.1016/j.ygyno.2016.10.031. URL: http://dx.doi.org/10.1016/j.ygyno.2016.10.031.

[81] Hans Ikenberg et al. "Screening for Cervical Cancer Precursors With p16 / Ki-67 Dual-Stained Cytology : Results of the PALMS Study". In: *JNCI* 105.20 (2013). DOI: 10.1093/jnci/djt235.

[82] Catherine Witney. *Primary HPV screening training resources launched*. 2019. URL: https://phescreening.blog.gov.uk/2019/02/11/primary-hpv-screening-training-resources-launched/.

[83] *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy*. Dover Publications Inc., 1978.

[84] Jeffrey L. Suhalim et al. "The need for speed". In: *Journal of Biophotonics* 5.5-6 (2012), pp. 387–395. ISSN: 15378276. DOI: 10.1002/jbio.201200002. arXiv: NIHMS150003. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf.

[85] Z. Nie, A. Petukhova, and E. Kumacheva. "Properties and emerging applications of self-assembled structures made from inorganic nanoparticles". In: *Nature Nanotechnology* 5 (Feb. 2010), pp. 15–25. DOI: 10.1038/nnano.2009.453.

[86] Andrew P. Shreve, Nerine J. Cherepy, and Richard A. Mathies. "Effective Rejection of Fluorescence Interference in Raman Spectroscopy Using a Shifted Excitation Difference Technique". In: *Applied Spectroscopy* 46.4 (1992), pp. 707–711. ISSN: 0003-7028. DOI: 10.1366/0003702924125122.

[87] Max Diem et al. "Applications of infrared and raman microspectroscopy of cells and tissue in medical diagnostics: Present status and future promises". In: *Spectroscopy (New York)* 27.5-6 (2012), pp. 463–496. ISSN: 07124813. DOI: 10.1155/2012/848360.

[88] Praveen D. Chowdary et al. "Molecular histopathology by spectrally reconstructed nonlinear interferometric vibrational imaging". In: *Cancer Research* 70.23 (2010), pp. 9562–9569. ISSN: 00085472. DOI: 10.1158/0008-5472.CAN-10-1554.

[89] Lauren A. Austin, Sam Osseiran, and Conor L. Evans. "Raman technologies in cancer diagnostics". In: *Analyst* 141.2 (2016), pp. 476–503. ISSN: 13645528. DOI: 10.1039/c5an01786f.

[90] Christoph Krafft et al. "Label-Free Molecular Imaging of Biological Cells and Tissues by Linear and Nonlinear Raman Spectroscopic Approaches". In: *Angewandte Chemie - International Edition* 56.16 (2017), pp. 4392–4430. ISSN: 15213773. DOI: 10.1002/anie.201607604.

[91] Kenny Kong et al. "Raman spectroscopy for medical diagnostics - From in-vitro biofluid assays to in-vivo cancer detection". In: *Advanced Drug Delivery Reviews* 89 (2015), pp. 121–134. ISSN: 18728294. DOI: 10.1016/j.addr.2015.03.009.

[92] Ji Xin Cheng and X. Sunney Xie. "Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine". In: *Science* 350.6264 (2015). ISSN: 10959203. DOI: 10.1126/science.aaa8870.

[93] Eva Brauchle et al. "Cell death stages in single apoptotic and necrotic cells monitored by Raman microspectroscopy". In: *Scientific Reports* 4 (2014), pp. 1–9. ISSN: 20452322. DOI: 10.1038/srep04698.

[94] Eva Brauchle et al. "Non-invasive Chamber-Specific Identification of Cardiomyocytes in Differentiating Pluripotent Stem Cells". In: *Stem Cell Reports* 6.2 (2016), pp. 188–199. ISSN: 22136711. DOI: 10.1016/j.stemcr.2015.12.007. URL: http://dx.doi.org/10.1016/j.stemcr.2015.12.007.

[95] Yasuaki Kumamoto et al. "Label-free molecular imaging and analysis by Raman spectroscopy". In: *Acta Histochemica et Cytochemica* 51.3 (2018), pp. 101–110. ISSN: 13475800. DOI: 10.1267/ahc.18019.

[96] Eliana Cordero. "In-vivo Raman spectroscopy: from basics to applications". In: *Journal of Biomedical Optics* 23.07 (2018), p. 1. ISSN: 1560-2281. DOI: 10.1117/1.jbo.23.7.071210.

[97] Miguel G. Ramírez-Elías and Francisco J. González. "Raman Spectroscopy for In Vivo Medical Diagnosis". In: *IntechOpen* (2018). DOI: http://dx.doi.org/10.5772/57353. URL: https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics.

[98] Nandan K. Das et al. "Raman plus X: Biomedical applications of multimodal Raman spectroscopy". In: *Sensors (Switzerland)* 17.7 (2017). ISSN: 14248220. DOI: 10.3390/s17071592.

[99] Sarah McAughtrie, Karen Faulds, and Duncan Graham. "Surface enhanced Raman spectroscopy (SERS): Potential applications for disease detection and treatment". In: *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* 21 (2014), pp. 40–53. ISSN: 13895567. DOI: 10.1016/j.jphotochemrev.2014.09.002. URL: http://dx.doi.org/10.1016/j.jphotochemrev.2014.09.002.

[100] S Shanmukh et al. "Identification and classification of respiratory syncytial virus (RSV) strains by surface-enhanced Raman spectroscopy and multivariate statistical techniques." In: *Analytical and bioanalytical chemistry* 390.6 (2008), pp. 1551–5. ISSN: 1618-2650. DOI: 10.1007/s00216-008-1851-0. URL: http://www.ncbi.nlm.nih.gov/pubmed/18236030.

[101] Vinh Hoang et al. "Identification of individual genotypes of measles virus using surface enhanced Raman spectroscopy". In: *Analyst* 135.12 (2010), pp. 3103–3109. ISSN: 00032654. DOI: 10.1039/c0an00453g.

[102] Jeremy D. Driskell et al. "Rapid and sensitive detection of rotavirus molecular signatures using surface enhanced raman spectroscopy". In: *PLoS ONE* 5.4 (2010). ISSN: 19326203. DOI: 10.1371/journal.pone.0010222.

[103] Janina Kneipp et al. "In vivo molecular probing of cellular compartments with gold nanoparticles and nanoaggregates". In: *Nano Letters* 6.10 (2006), pp. 2225–2231. ISSN: 15306984. DOI: 10.1021/nl061517x.

[104] Shawn P. Mulvaney et al. "Three-layer substrates for surface-enhanced Raman scattering: Preparation and preliminary evaluation". In: *Journal of Raman Spectroscopy* 34.2 (2003), pp. 163–171. ISSN: 03770486. DOI: 10.1002/jrs.972.

[105] Rebecca S. Golightly, William E. Doering, and Michael J. Natan. "Surface-enhanced Raman spectroscopy and homeland security: A perfect match?" In: *ACS Nano* 3.10 (2009), pp. 2859–2869. ISSN: 19360851. DOI: 10.1021/nn9013593.

[106] S. Efrima and L. Zeiri. "Understanding SERS of bacteria". In: *Journal of Raman Spectroscopy* 40.3 (2009), pp. 277–288. ISSN: 03770486. DOI: 10.1002/jrs.2121.

[107] I-hsien Chou et al. "Nanofluidic Biosensing for β-amyloid Detection Using Surface Enhanced Raman Spectroscopy (SERS)". In: *Nano Letters* 8.6 (2008), pp. 1729–1735. DOI: 10.1021/nl0808132.Nanofluidic.

[108] Alessia Pallaoro et al. "Rapid identification by surface-enhanced raman spectroscopy of cancer cells at low concentrations flowing in a microfluidic channel". In: *ACS Nano* 9.4 (2015), pp. 4328–4336. ISSN: 1936086X. DOI: 10.1021/acsnano.5b00750.

[109] Kazuki Bando et al. "Analysis of dynamic SERS spectra measured with a nanoparticle during intracellular transportation in 3D". In: *Journal of Optics (United Kingdom)* 17.11 (2015). ISSN: 20408986. DOI: 10.1088/2040-8978/17/11/114023.

[110] Anna Huefner et al. "Characterization and visualization of vesicles in the endolysosomal pathway with surface-enhanced raman spectroscopy and chemometrics". In: *ACS Nano* 10.1 (2016), pp. 307–316. ISSN: 1936086X. DOI: 10.1021/acsnano.5b04456.

[111] Jing Yang et al. "Distinguishing breast cancer cells using surface-enhanced Raman scattering". In: *Analytical and Bioanalytical Chemistry* 402.3 (2012), pp. 1093–1100. ISSN: 16182642. DOI: 10.1007/s00216-011-5577-z.

[112] Gufeng Wang et al. "Detection of the Potential Pancreatic Cancer Marker MUC4 in Serum Using Surface-Enhanced Raman Scattering". In: *Analytical Chemistry* 83.7 (2011), pp. 2554–2561. ISSN: 15378276. DOI: 10.1021/ac102829b. arXiv: NIHMS150003.

[113] Anita Mahadevan-Jansen et al. "Development of a Fiber Optic Probe to Measure NIR Raman Spectra of Cervical Tissue in Vivo". In: *Photochemistry and Photobiology* 68.3 (1998), pp. 427–431. ISSN: 00318655. DOI: 10.1111/j.1751-1097.1998.tb09703.x.

[114] U. Utzinger et al. "Near-infrared Raman spectroscopy for in vivo detection of cervical precancers". In: *Applied Spectroscopy* 55.8 (2001), pp. 955–959. ISSN: 00037028. DOI: 10.1366/0003702011953018.

[115] Kamila Magdalena Ostrowska et al. "Investigation of the influence of high-risk human papillomavirus on the biochemical composition of cervical cancer cells using vibrational spectroscopy." In: *The Analyst* 135.12 (2010), pp. 3087–93. ISSN: 1364-5528. DOI: 10.1039/c0an00571a. URL: http://www.ncbi.nlm.nih.gov/pubmed/20967345.

[116] Youseph Yazdi et al. "Resonance Raman Spectroscopy at 257nm Excitation of Normal and Malignant Cultured Brest and Cervical Cells". In: *Applied Spectroscopy* 53.1 (1999), pp. 82–85.

[117] Philip R T Jess et al. "Early detection of cervical neoplasia by Raman spectroscopy". In: *International Journal of Cancer* 121.12 (2007), pp. 2723–2728. ISSN: 00207136. DOI: 10.1002/ijc.23046.

[118] Dong Hyun Kim et al. "Raman chemical mapping reveals site of action of HIV protease inhibitors in HPV16 E6 expressing cervical carcinoma cells". In: *Analytical and Bioanalytical Chemistry* 398.7-8 (2010), pp. 3051–3061. ISSN: 16182642. DOI: 10.1007/s00216-010-4283-6.

[119] S. Rubina et al. "Raman spectroscopic study on classification of cervical cell specimens". In: *Vibrational Spectroscopy* 68 (2013), pp. 115–121. ISSN: 09242031. DOI: 10.1016/j.vibspec.2013.06.002. URL: http://dx.doi.org/10.1016/j.vibspec.2013.06.002.

[120] Elizabeth M. Kanter et al. "Effect of hormonal variation on Raman spectra for cervical disease detection". In: *Arn J Obstet Gynecol.* 200.5 (2010), pp. 1–13. DOI: 10.1016/j.ajog.2008.11.024.Effect.

[121] Elizabeth Vargis et al. "Effect of normal variations on disease classification of Raman spectra from cervical tissue". In: *The Analyst* 136 (2011), pp. 2981–2987. ISSN: 0003-2654. DOI: 10.1039/c0an01020k.

[122] Shiyamala Duraipandian et al. "In vivo diagnosis of cervical precancer using Raman spectroscopy and genetic algorithm techniques". In: *Analyst* 136.20 (2011), pp. 4328–4336. ISSN: 00032654. DOI: 10.1039/c1an15296c.

[123] Shiyamala Duraipandian et al. "Simultaneous fingerprint and high-wavenumber confocal Raman spectroscopy enhances early detection of cervical precancer in vivo". In: *Analytical Chemistry* 84.14 (2012), pp. 5913–5919. ISSN: 00032700. DOI: 10.1021/ac300394f.

[124] Shangyuan Feng et al. "Blood plasma surface-enhanced Raman spectroscopy for non-invasive optical detection of cervical cancer." In: *The Analyst* 138.14 (2013), pp. 3967–74. ISSN: 1364-5528. DOI: 10.1039/c3an36890d. URL: http://www.ncbi.nlm.nih.gov/pubmed/23529624.

[125] Nisha Narayanan et al. "Aggregation induced Raman scattering of squaraine dye: Implementation in diagnosis of cervical cancer dysplasia by SERS imaging". In: *Biosensors and Bioelectronics* 70 (2015), pp. 145–152. ISSN: 18734235. DOI: 10.1016/j.bios.2015.03.029. URL: http://dx.doi.org/10.1016/j.bios.2015.03.029.

[126] Ernesto Guccione et al. "Comparative analysis of the intracellular location of the high- and low-risk human papillomavirus oncoproteins". In: *Virology* 293.1 (2002), pp. 20–25. ISSN: 00426822. DOI: 10.1006/viro.2001.1290.

[127] Dong Wei, Shuo Chen, and Quan Liu. "Review of fluorescence suppression techniques in Raman spectroscopy". In: *Applied Spectroscopy Reviews* 50.5 (2015), pp. 387–406. ISSN: 1520569X. DOI: 10.1080/05704928.2014.999936.

[128] P. Matousek et al. "Efficient rejection of fluorescence from Raman spectra using picosecond Kerr gating". In: *Applied Spectroscopy* 53.12 (1999), pp. 1485–1489. ISSN: 00037028. DOI: 10.1366/0003702991945993.

[129] D. V. Martyshkin et al. "Effective suppression of fluorescence light in Raman measurements using ultrafast time gated charge coupled device camera". In: *Review of Scientific Instruments* 75.3 (2004), pp. 630–635. ISSN: 0034-6748. DOI: 10.1063/1.1646743.

[130] Juha Kostamovaara et al. "Fluorescence suppression in Raman spectroscopy using a time-gated CMOS SPAD". In: *Optics Express* 21.25 (2013), p. 31632. ISSN: 1094-4087. DOI: 10.1364/OE.21.031632. URL: https://www.osapublishing.org/oe/abstract.cfm?uri=oe-21-25-31632.

[131] T. Tahara and H.-O. Hamaguchi. "Picosecond Raman Spectroscopy Using a Streak Camera". In: *Science And Technology* 47.4 (1993).

[132] W B Spillman et al. "A New Technique for the Elimination of Fluorescence Interference in Raman Spectroscopy". In: *Rapid communications* 40.5 (1988), pp. 1–4.

[133] M. J. Wirth and Shiow Hwa Chou. "Comparison of Time and Frequency Domain Methods for Rejecting Fluorescence from Raman Spectra". In: *Analytical Chemistry* 60.18 (1988), pp. 1882–1886. ISSN: 15206882. DOI: 10.1021/ac00169a009.

[134] John B. Cooper, Mohamed Abdelkader, and Kent L. Wise. "Sequentially shifted excitation Raman spectroscopy: Novel algorithm and instrumentation for fluorescence-free Raman spectroscopy in spectral space". In: *Applied Spectroscopy* 67.8 (2013), pp. 973–984. ISSN: 00037028. DOI: 10.1366/12-06852.

[135] Francesca Rosi et al. "Subtracted shifted Raman spectroscopy of organic dyes and lakes". In: *Journal of Raman Spectroscopy* 41.4 (2010), pp. 452–458. ISSN: 03770486. DOI: 10.1002/jrs.2447.

[136] Anna Chiara De Luca et al. "Online fluorescence suppression in modulated raman spectroscopy". In: *Analytical Chemistry* 82.2 (2010), pp. 738–745. ISSN: 00032700. DOI: 10.1021/ac9026737.

[137] Mingzhou Chen et al. "The use of wavelength modulated raman spectroscopy in label-free identification of T lymphocyte subsets, natural killer cells and dendritic Cells". In: *PLoS ONE* 10.5 (2015), pp. 1–14. ISSN: 19326203. DOI: 10.1371/journal.pone.0125158.

[138] Bavishna B. Praveen et al. "Optimisation of Wavelength Modulated Raman Spectroscopy: Towards High Throughput Cell Screening". In: *PLoS ONE* 8.6 (2013), pp. 4–8. ISSN: 19326203. DOI: 10.1371/journal.pone.0067211.

[139] Elisabetta Canetta et al. "Modulated Raman spectroscopy for enhanced identification of bladder tumor cells in urine samples." In: *Journal of biomedical optics* 16 (2011), p. 037002. ISSN: 10833668. DOI: 10.1117/1.3556722.

[140] Michael Mazilu et al. "Optimal algorithm for fluorescence suppression of modulated Raman spectroscopy Abstract :" in: *Optics Express* 18.11 (2010), pp. 1082–1088. ISSN: 1094-4087. DOI: 10.1364/OE.18.011382.

[141] B M Wise et al. "Theoretical basis for the use of principal component models for monitoring multivariate processes". In: *Process Control and Quality* 1.1 (1990), pp. 41–51. ISSN: 09243089 (ISSN).

[142] Herve Abdi and Lynne J Williams. "Principal component analysis". In: *WIREs Computational Statistics* (2010), pp. 374–377. DOI: 10.1016/B978-0-08-044894-7.01358-0.

[143] Ramya Bhatia. *The Scottish HPV Archive (Report)*. 2017. URL: http://www.ed.ac.uk/pathology/research/scottish-hpv-archive.

[144] Alison J. Hobro and Nicholas I. Smith. "An evaluation of fixation methods: Spatial and compositional cellular changes observed by Raman imaging". In: *Vibrational Spectroscopy* 91 (2017), pp. 31–45. ISSN: 09242031. DOI: 10.1016/j.vibspec.2016.10.012. URL: http://dx.doi.org/10.1016/j.vibspec.2016.10.012.

[145] Mingzhou Chen et al. "Depth-resolved multimodal imaging: Wavelength modulated spatially offset Raman spectroscopy with optical coherence tomography". In: *Journal of Biophotonics* 11.1 (2018), pp. 1–7. ISSN: 18640648. DOI: 10.1002/jbio.201700129.

[146] Naomi McReynolds et al. "Multimodal discrimination of immune cells using a combination of Raman spectroscopy and digital holographic microscopy". In: *Scientific Reports* 7.September 2016 (2017), pp. 1–11. ISSN: 20452322. DOI: 10.1038/srep43631. URL: http://dx.doi.org/10.1038/srep43631.

[147] Xiao Shan Zheng et al. "Label-free SERS in biological and biomedical applications: Recent progress, current challenges and opportunities". In: *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* 197 (2018), pp. 56–77. ISSN: 13861425. DOI: 10.1016/j.saa.2018.01.063. URL: https://doi.org/10.1016/j.saa.2018.01.063.

[148] Xiaotian Wang et al. "Surface-Enhanced Raman Scattering (SERS) on transition metal and semiconductor nanostructures". In: *Physical Chemistry Chemical Physics* 14.17 (2012), pp. 5891–5901. ISSN: 14639076. DOI: 10.1039/c2cp40080d.

[149] Nguyen T K Thanh and Luke A W Green. "Functionalisation of nanoparticles for biomedical applications". In: *Nano Today* 5.3 (2010), pp. 213–230. ISSN: 17480132. DOI: 10.1016/j.nantod.2010.05.003.

[150] Gokhan Demirel et al. "Surface-enhanced Raman spectroscopy (SERS): An adventure from plasmonic metals to organic semiconductors as SERS platforms". In: *Journal of Materials Chemistry C* 6.20 (2018), pp. 5314–5335. ISSN: 20507526. DOI: 10.1039/c8tc01168k.

[151] Catherine J. Murphy et al. "ChemInform Abstract: Gold Nanoparticles in Biology: Beyond Toxicity to Cellular Imaging". In: *ChemInform* 40.16 (2009). ISSN: 09317597. DOI: 10.1002/chin.200916266.

[152] Alaaldin M. Alkilany and Catherine J. Murphy. "Toxicity and cellular uptake of gold nanoparticles: What we have learned so far?" In: *Journal of Nanoparticle Research* 12.7 (2010), pp. 2313–2333. ISSN: 13880764. DOI: 10.1007/s11051-010-9911-8.

[153] J. Kimling et al. "Turkevich method for gold nanoparticle synthesis revisited". In: *Journal of Physical Chemistry B* 110.32 (2006), pp. 15700–15707. ISSN: 15206106. DOI: 10.1021/jp061667w.

[154] Roya Herizchi et al. "Current methods for synthesis of gold nanoparticles". In: *Artificial Cells, Nanomedicine and Biotechnology* 44.2 (2016), pp. 596–602. ISSN: 2169141X. DOI: 10.3109/21691401.2014.971807.

[155] Pierre Colson, Catherine Henrist, and Rudi Cloots. "Nanosphere Lithography: A Powerful Method for the Controlled Manufacturing of Nanomaterials". In: *Journal of Nanomaterials* 2013 (2013), pp. 1–19. ISSN: 1687-4110. DOI: 10.1155/2013/948510.

[156] Hai Le-The et al. "Large-scale fabrication of highly ordered sub-20 nm noble metal nanoparticles on silica substrates without metallic adhesion layers". In: *Microsystems & Nanoengineering* 4.1 (2018), pp. 1–10. ISSN: 2055-7434. DOI: 10.1038/s41378-017-0001-2. URL: http://dx.doi.org/10.1038/s41378-017-0001-2.

[157] M. Vinod and K. G. Gopchandran. "Au, Ag and Au: Ag colloidal nanoparticles synthesized by pulsed laser ablation as SERS substrates". In: *Progress in Natural Science: Materials International* 24.6 (2014), pp. 569–578. ISSN: 10020071. DOI: 10.1016/j.pnsc.2014.10.003.

[158] Gloria Herrera, Amira Padilla, and Samuel Hernandez-Rivera. "Surface Enhanced Raman Scattering (SERS) Studies of Gold and Silver Nanoparticles Prepared by Laser Ablation". In: *Nanomaterials* 3.1 (2013), pp. 158–172. DOI: 10.3390/nano3010158.

[159] K. Vijayaraghavan and T. Ashokkumar. "Plant-mediated biosynthesis of metallic nanoparticles: A review of literature, factors affecting synthesis, characterization techniques and applications". In: *Journal of Environmental Chemical Engineering* 5.5 (2017), pp. 4866–4883. ISSN: 22133437. DOI: 10.1016/j.jece.2017.09.026. URL: http://dx.doi.org/10.1016/j.jece.2017.09.026.

[160] Muhammad Ovais et al. "Biosynthesis of metal nanoparticles via microbial enzymes: A mechanistic approach". In: *International Journal of Molecular Sciences* 19.12 (2018), pp. 1–20. ISSN: 14220067. DOI: 10.3390/ijms19124100.

[161] José M. Romo-Herrera, Ramón A. Alvarez-Puebla, and Luis M. Liz-Marzán. "Controlled assembly of plasmonic colloidal nanoparticle clusters". In: *Nanoscale* 3.4 (2011), pp. 1304–1315. ISSN: 20403364. DOI: 10.1039/c0nr00804d.

[162] Bernd Küstner et al. "SERS labels for red laser excitation: Silica-encapsulated SAMs on tunable gold/silver nanoshells". In: *Angewandte Chemie - International Edition* 48.11 (2009), pp. 1950–1953. ISSN: 14337851. DOI: 10.1002/anie.200804518.

[163] Rajib Ghosh Chaudhuri and Santanu Paria. "Core/Shell Nanoparticles: Classes, Properties, Synthesis Mechanisms, Characterization, and Applications". In: *Chemical Reviews* 112.4 (2012), pp. 2373–2433. ISSN: 0009-2665. DOI: 10.1021/cr100449n.

[164] Bong Hyun Jun et al. "Multifunctional silver-embedded magnetic nanoparticles as SERS nanoprobes and their applications". In: *Small* 6.1 (2010), pp. 119–125. ISSN: 16136810. DOI: 10.1002/smll.200901459.

[165] Pamela Mosier-Boss. "Review of SERS Substrates for Chemical Sensing". In: *Nanomaterials* 7.6 (2017), p. 142. ISSN: 2079-4991. DOI: 10.3390/nano7060142. URL: http://www.mdpi.com/2079-4991/7/6/142.

[166] George C. Schatz et al. "Wavelength-scanned surface-enhanced resonance raman excitation spectroscopy". In: *Journal of Physical Chemistry C* 112.49 (2008), pp. 19302–19310. ISSN: 19327447. DOI: 10.1021/jp807837t.

[167] Yuling Wang and Sebastian Schlücker. "Rational design and synthesis of SERS labels". In: *Analyst* 138.8 (2013), pp. 2224–2238. ISSN: 13645528. DOI: 10.1039/c3an36866a.

[168] Fanghao Hu et al. "Live-Cell Bioorthogonal Chemical Imaging: Stimulated Raman Scattering Microscopy of Vibrational Probes". In: *Accounts of Chemical Research* 49.8 (2016), pp. 1494–1502. ISSN: 0001-4842. DOI: 10.1021/acs.accounts.6b00210.

[169] Christopher D. Spicer et al. "Peptide and protein nanoparticle conjugates: Versatile platforms for biomedical applications". In: *Chemical Society Reviews* 47.10 (2018), pp. 3574–3620. ISSN: 14604744. DOI: 10.1039/c7cs00877e. URL: http://dx.doi.org/10.1039/C7CS00877E.

[170] André F. Moreira et al. "Gold-core silica shell nanoparticles application in imaging and therapy: A review". In: *Microporous and Mesoporous Materials* 270.April (2018), pp. 168–179. ISSN: 13871811. DOI: 10.1016/j.micromeso.2018.05.022. URL: https://doi.org/10.1016/j.micromeso.2018.05.022.

[171] Luca Guerrini, Ramon A. Alvarez-Puebla, and Nicolas Pazos-Perez. "Surface modifications of nanoparticles for stability in biological fluids". In: *Materials* 11.7 (2018), pp. 1–28. ISSN: 19961944. DOI: 10.3390/ma11071154.

[172] Duncan Graham et al. "Control of enhanced Raman scattering using a DNA-based assembly process of dye-coded nanoparticles". In: *Nature Nanotechnology* 3.9 (2008), pp. 548–551. ISSN: 17483395. DOI: 10.1038/nnano.2008.189.

[173] Anirban Samanta and Igor L. Medintz. "Nanoparticles and DNA-a powerful and growing functional combination in bionanotechnology". In: *Nanoscale* 8.17 (2016), pp. 9037–9095. ISSN: 20403372. DOI: 10.1039/c5nr08465b.

[174] Chun Chi Wang et al. "Biomedical Applications of DNA-Conjugated Gold Nanoparticles". In: *ChemBioChem* (2016), pp. 1052–1062. ISSN: 14397633. DOI: 10.1002/cbic.201600014.

[175] Álvaro Artiga et al. "Current status and future perspectives of gold nanoparticle vectors for siRNA delivery". In: *Journal of Materials Chemistry B* 7.6 (2019), pp. 876–896. ISSN: 2050750X. DOI: 10.1039/c8tb02484g.

[176] Manuel Arruebo, Mónica Valladares, and África González-Fernández. "Antibody-conjugated nanoparticles for biomedical applications". In: *Journal of Nanomaterials* 2009 (2009). ISSN: 16874110. DOI: 10.1155/2009/439389.

[177] Daniel A Richards, Antoine Maruani, and Vijay Chudasama. "Chemical Science Antibody fragments as nanoparticle targeting ligands : a step in the right direction". In: *Chemical Science* 00 (2016), pp. 1–15. ISSN: 2041-6520. DOI: 10.1039/C6SC02403C. URL: http://dx.doi.org/10.1039/C6SC02403C.

[178] Abbas Alibakhshi et al. "Targeted cancer therapy through antibody fragments-decorated nanomedicines". In: *Journal of Controlled Release* 268.October (2017), pp. 323–334. ISSN: 18734995. DOI: 10.1016/j.jconrel.2017.10.036. URL: https://doi.org/10.1016/j.jconrel.2017.10.036.

[179] W. Kemner. "Separation of tumor cells from a suspension of dissociated human colorectal carcinoma tissue by means of monoclonal antibody-coated magnetic beads". In: *Journal of immunological methods.* 147.2 (1992), pp. 197–200. ISSN: 0022-1759.

[180] S Ozawa et al. "What's new in imaging? New magnetic resonance imaging of esophageal cancer using an endoluminal surface coil and antibody-coated magnetite particles". In: *Recent results in cancer research. Fortschritte der Krebsforschung. Progrès dans les recherches sur le cancer* 155 (Feb. 2000), pp. 73–87. DOI: 10.1007/978-3-642-59600-1_7.

[181] Mílton Cordeiro et al. "Gold nanoparticles for diagnostics: Advances towards points of care". In: *Diagnostics* 6.4 (2016). ISSN: 20754418. DOI: 10.3390/diagnostics6040043.

[182] Dvir Yelin et al. "Multiphoton plasmon-resonance microscopy". In: *Optics express* 11.12 (2003), pp. 1385–1391.

[183] C.B. Raub, E.J. Orwin, and R. Haskell. "Immunogold labeling to enhance contrast in optical coherence microscopy of tissue engineered corneal constructs". In: *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 1 (2005), pp. 1210–1213. DOI: 10.1109/iembs.2004.1403386.

[184] Xiong Liu et al. "A one-step homogeneous immunoassay for cancer biomarker detection using gold nanoparticle probes coupled with dynamic light scattering". In: *Journal of the American Chemical Society* 130.9 (2008), pp. 2780–2782. ISSN: 00027863. DOI: 10.1021/ja711298b.

[185] Ivan H. El-Sayed, Xiaohua Huang, and Mostafa A. El-Sayed. "Surface plasmon resonance scattering and absorption of anti-EGFR antibody conjugated gold nanoparticles in cancer diagnostics: Applications in oral cancer". In: *Nano Letters* 5.5 (2005), pp. 829–834. ISSN: 15306984. DOI: 10.1021/nl050074e.

[186] Pattasuda Duangkaew et al. "Biosensors and Bioelectronics Ultrasensitive electrochemical immunosensor based on dual signal ampli fi cation process for p16 INK4a cervical cancer detection in clinical samples". In: *Biosensors and Bioelectronic* 74 (2015), pp. 673–679. ISSN: 0956-5663. DOI: 10.1016/j.bios.2015.07.004. URL: http://dx.doi.org/10.1016/j.bios.2015.07.004.

[187] Ruchuon Yokchom, Somsak Laiwejpithaya, and Weerakanya Maneeprakorn. "Paper-based immunosensor with signal amplification by enzyme-labeled anti-p16 INK4a multifunctionalized gold nanoparticles for cervical cancer screening". In: *Nanomedicine: Nanotechnology, Biology, and Medicine* 14.3 (2018), pp. 1051–1058. ISSN: 1549-9634. DOI: 10.1016/j.nano.2018.01.016. URL: https://doi.org/10.1016/j.nano.2018.01.016.

[188] Greg T. Hermanson. "Microparticles and Nanoparticles". In: *Bioconjugate Techniques*. 2013, pp. 360–388. ISBN: 9780123746269. DOI: 10.1016/B978-0-08-087780-8.00034-6.

[189] Greg T. Hermanson. *Immobilization of Ligands on Chromatography Supports*. 2013, pp. 589–740. ISBN: 9780123822390. DOI: 10.1016/b978-0-12-382239-0.00015-7.

[190] Mir Hadi Jazayeri et al. "Various methods of gold nanoparticles (GNPs) conjugation to antibodies". In: *Sensing and Bio-Sensing Research* 9 (2016), pp. 17–22. ISSN: 22141804. DOI: 10.1016/j.sbsr.2016.04.002. URL: http://dx.doi.org/10.1016/j.sbsr.2016.04.002.

[191] Mehmet Kahraman et al. "Fundamentals and applications of SERS-based bioanalytical sensing". In: *Nanophotonics* 6.5 (2017), pp. 831–852. DOI: 10.1515/nanoph-2016-0174.

[192] *Principles of Surface-Enhanced Raman Spectroscopy*. Elsevier, 2009.

[193] Christy Haynes, Adam D Mcfarland, and Richard P. Van Duyne. "Surface-enhanced Raman spectroscopy". In: *Analytical Chemistry* (2005). ISSN: 09718044. DOI: 10.1007/s12045-010-0016-6.

[194] Luca Guerrini and Duncan Graham. "Molecularly-mediated assemblies of plasmonic nanoparticles for Surface-Enhanced Raman Spectroscopy applications". In: *Chemical Society Reviews* 41.21 (2012), pp. 7085–7107. ISSN: 03060012. DOI: 10.1039/c2cs35118h.

[195] Wolfgang Haiss et al. "Determination of size and concentration of gold nanoparticles from UV-Vis spectra". In: *Analytical Chemistry* 79.11 (2007), pp. 4215–4221. ISSN: 00032700. DOI: 10.1021/ac0702084.

[196] David A Armbruster and Terry Pry. "Limit of blank, limit of detection and limit of quantitation." In: *The Clinical biochemist. Reviews* 29 Suppl 1.August (2008), S49–52. ISSN: 0159-8090. URL: http://www.ncbi.nlm.nih.gov/pubmed/18852857{\%}0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2556583.

[197] Maria Navas-Moreno et al. "Nanoparticles for live cell microscopy: A surface-enhanced Raman scattering perspective". In: *Scientific Reports* 7.1 (2017), pp. 3–9. ISSN: 20452322. DOI: 10.1038/s41598-017-04066-0.

[198] Mine Altunbek, Gamze Kuku, and Mustafa Culha. "Gold nanoparticles in single-cell analysis for surface enhanced Raman scattering". In: *Molecules* 21.12 (2016), pp. 1–18. ISSN: 14203049. DOI: 10.3390/molecules21121617.

[199] John P Nolan et al. "Single cell analysis using surface enhanced Raman scattering (SERS) tags". In: *Methods* 57.3 (2013), pp. 272–279. DOI: 10.1016/j.ymeth.2012.03.024.Single.

[200] Suwussa Bamrungsap et al. "SERS-fluorescence dual mode nanotags for cervical cancer detection using aptamers conjugated to gold-silver nanorods". In: *Microchimica Acta* 183.1 (2016), pp. 249–256. ISSN: 0026-3672. DOI: 10.1007/s00604-015-1639-9. URL: http://link.springer.com/10.1007/s00604-015-1639-9.

[201] Jun-jie Sun et al. "The Increased PTK7 Expression Is a Malignant Factor in Cervical Cancer". In: *Disease markers* 2019 (2019).

[202] Zhilun Zhao et al. "Applications of vibrational tags in biological imaging by Raman microscopy". In: *The Analyst* (2017), pp. 4018–4029. ISSN: 0003-2654. DOI: 10.1039/C7AN01001J. URL: http://pubs.rsc.org/en/Content/ArticleLanding/2017/AN/C7AN01001J.

[203] Luis M Campos et al. "Stimulated Raman scattering of polymer nanoparticles for multiplexed live-cell imaging". In: *Chemical communications* 53.46 (2017). DOI: 10.1039/c7cc01860f.

[204] W. J. Tipping et al. "Stimulated Raman scattering microscopy: an emerging tool for drug discovery". In: *Chem. Soc. Rev.* 45.8 (2016), pp. 2075–2089. ISSN: 0306-0012. DOI: 10.1039/C5CS00693G. URL: http://xlink.rsc.org/?DOI=C5CS00693G.

[205] Fanghao Hu, Yihui Shen, and Lu Wei. "Vibrational imaging of glucose uptake activity in live cells and tissues by stimulated Raman scattering". In: *Angewandte Chemie - International Edition* 54.34 (2016), pp. 9821–9825. DOI: 10.1002/anie.201502543.Vibrational.

[206] Yu Winston Wang et al. "Rapid ratiometric biomarker detection with topically applied SERS nanoparticles." In: *Technology* 2.2 (2014), pp. 118–132. ISSN: 2339-5478. DOI: 10.1142/S2339547814500125. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4103661{\&}tool=pmcentrez{\&}rendertype=abstract.

[207] Moritz F Kircher et al. "A brain tumor molecular imaging strategy using a new triple-modality MRI-photoacoustic-Raman nanoparticle". In: *Nature Medicine* 18.5 (2012), pp. 829–835. ISSN: 1078-8956. DOI: 10.1038/nm.2721.

[208] Jonathan T C Liu et al. "Quantifying Cell-Surface Biomarker Expression in Thick Tissues with Ratiometric Three-Dimensional Microscopy". In: *Biophysj* 96.6 (2009), pp. 2405–2414. ISSN: 0006-3495. DOI: 10.1016/j.bpj.2008.12.3908. URL: http://dx.doi.org/10.1016/j.bpj.2008.12.3908.

[209] Ana Pérez-Jiménez et al. "Surface-enhanced Raman spectroscopy: benefits, trade-offs and future developments". In: *Chem. Sci.* (2020).

[210] Steven Y. L. Leigh, Madhura Som, and Jonathan T.C. Liu. "Method for Assessing the Reliability of Molecular Diagnostics Based on Multiplexed SERS-Coded Nanoparticles". In: *PLOS One* (2013).

[211] Cristina L. Zavaleta et al. "Multiplexed imaging of surface enhanced Raman scattering nanotags in living mice using noninvasive Raman spectroscopy". In: *PNAS* (2009).

[212] William J Tipping et al. "Imaging drug uptake by bioorthogonal stimulated Raman scattering microscopy". In: *Chemical Science* (2017), pp. 5606–5615. DOI: 10.1039/c7sc01837a.

[213] Bengt Fadeel. "Hide and Seek : Nanomaterial Interactions With the Immune System". In: *Frontiers in Immunology* 10.February (2019), pp. 1–10. DOI: 10.3389/fimmu.2019.00133.

[214] Daniel F Moyano et al. "Immunomodulatory Effects of Coated Gold Nanoparticles in LPS-Stimulated In Vitro and In Vivo Murine Model Systems Immunomodulatory Effects of Coated Gold Nanoparticles in LPS-Stimulated In Vitro and In Vivo Murine Model Systems". In: *Chem* (2016), pp. 320–327. DOI: 10.1016/j.chempr.2016.07.007.

[215] Giulio Caracciolo et al. "Stealth Effect of Biomolecular Corona on Nanoparticle Uptake by Immune Cells". In: *Langmuir* (2015). DOI: 10.1021/acs.langmuir.5b02158.

[216] Olimpia Gamucci et al. "Biomedical Nanoparticles: Overview of Their Surface Immune-Compatibility". In: *Coatings* (2014), pp. 139–159. DOI: 10.3390/coatings4010139.

[217] Laurence A Cole. "Pregnancy Tests". In: *Human Chorionic Gonadotropin (hCG)*. First Edit. Elsevier Inc., 2010, pp. 251–257. ISBN: 9780123849076. DOI: 10.1016/B978-0-12-384907-6.00023-2. URL: http://dx.doi.org/10.1016/B978-0-12-384907-6.00023-2.

[218] B Y A J P Martin and A D R L M Synge. "I5I . A NEW FORM OF CHROMATOGRAM EMPLOYING TWO LIQUID PHASES 2 . APPLICATION TO THE MICRO-DETERMINATION OF THE HIGHER MONOAMINO-ACIDS IN PROTEINS". In: *Biochemical Journal* (1941).

[219] By Rosalyn S Yalow and Solomon A Berson. "Immunoassay of Endogenous Plasma Insulin in Man". In: *Journal of clinical investigation* (1960), pp. 11–13.

[220] "Lateral Flow Immunoassay". In: Humana Press, 2009. Chap. Market Trends in Lateral Flow Immunoassays.

[221] Suzanne Smith et al. "The potential of paper-based diagnostics to meet the ASSURED criteria". In: *Rsc Advances* (2018), pp. 34012–34034. DOI: 10.1039/c8ra06132g.

[222] Duo Ma et al. "Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment". In: *Synthetic Biology* 3.September (2018), pp. 1–11. DOI: 10.1093/synbio/ysy018.

[223] Nili Ostrov et al. "A modular yeast biosensor for low-cost point-of-care pathogen detection". In: *Science Advances* (2017).

[224] Despina P Kalogianni et al. "Dry reagent dipstick test combined with 23S rRNA PCR for molecular diagnosis of bacterial infection in arthroplasty". In: *Analytical Biochemistry* 361 (2007), pp. 169–175. DOI: 10.1016/j.ab.2006.11.013.

[225] Elena A Zvereva et al. "Analytical Methods Cut-off on demand : adjustment of the threshold level of an immunochromatographic assay for chloramphenicol". In: *Analytical methods* (2015), pp. 6378–6384. DOI: 10.1039/c5ay00835b.

[226] Suquan Song et al. "Multiplex Lateral Flow Immunoassay for Mycotoxin Determination". In: *Analytical Chemistry* (2014). DOI: 10.1021/ac500540z.

[227] Adrian Carrio et al. "Automated Low-Cost Smartphone-Based Lateral Flow Saliva Test Reader for Drugs-of-Abuse Detection". In: *Sensors* (2015), pp. 29569–29593. DOI: 10.3390/s151129569.

[228] N De Giovanni and N Fucci. "The Current Status of Sweat Testing For Drugs of Abuse : A Review". In: *Current Medicinal Chemistry* (2013), pp. 545–561.

[229] Xiaonan Liu et al. "Multiple SNPs Detection Based on Lateral Flow Assay for Phenylketonuria Diagnostic". In: *Analytical Chemistry* (2018). DOI: 10.1021/acs.analchem.7b05113.

[230] Miriam Jauset-rubio et al. "Ultrasensitive , rapid and inexpensive detection of DNA using paper based lateral flow assay". In: *Nature Publishing Group* October (2016), pp. 1–10. DOI: 10.1038/srep37732. URL: http://dx.doi.org/10.1038/srep37732.

[231] Georgina M S Ross, Monique G E G Bremer, and Michel W F Nielen. "Consumer-friendly food allergen detection : moving towards smartphone-based immunoassays". In: *Analytical and Bioanalytical Chemistry* (2018), pp. 5353–5371.

[232] Song Cheng et al. "Fluorescent microspheres lateral flow assay for sensitive detection of the milk allergen casein". In: *Food and Agricultural Immunology* 28.6 (2017), pp. 1017–1028. ISSN: 0954-0105. DOI: 10.1080/09540105.2017.1325841. URL: https://doi.org/10.1080/09540105.2017.1325841.

[233] Fumiko Endo et al. "Development of a simple and quick immunochromatography method for detection of anti-HPV-16 / -18 antibodies". In: *PLoS ONE* (2017). DOI: 10.1371/journal.pone.0171314.

[234] Sammartino Mariano et al. "A Low-Cost HPV Immunochromatographic Assay to Detect High-Grade Cervical Intraepithelial Neoplasia". In: *PLoS ONE* (2016), pp. 1–12. DOI: 10.1371/journal.pone.0164892.

[235] Parvapan Bhattarakosol, Kamolwan Plaignam, and Amornpun Sereemaspun. "Immunogold-agglutination assay for direct detection of HPV-16 E6 and L1 proteins from clinical specimens". In: *Journal of Virological Methods* 255.May 2017 (2018), pp. 60–65. ISSN: 0166-0934. DOI: 10.1016/j.jviromet.2018.02.009. URL: https://doi.org/10.1016/j.jviromet.2018.02.009.

[236] Amrei Krings et al. "Performance of OncoE6 cervical test with collection methods enabling self-sampling". In: *BMC Women's Health* (2018), pp. 1–7.

[237] Natalia M. Rodriguez et al. "A Fully Integrated Paperfluidic Molecular Diagnostic Chip for the Extraction, Amplification, and Detection of Nucleic Acids from Clinical Samples". In: *Lab on a chip* 16.4 (2016), pp. 753–763. DOI: 10.1039/c5lc01392e.A.

[238] Cara S Kosack, Anne-laure Page, and Paul R Klatser. "A guide to aid the selection of diagnostic tests". In: *Bull World Health Organ* October 2016 (2017), pp. 639–645.

[239] Sherine F Cheung, Samantha K L Cheng, and Daniel T Kamei. "Paper-Based Systems for Point-of-Care Biosensing". In: *Journal of Laboratory Automation* (2015). DOI: 10.1177/2211068215577197.

[240] Cody Carrell et al. "Microelectronic Engineering Beyond the lateral fl ow assay : A review of paper-based microfluidics". In: *Microelectronic Engineering* 206.November 2018 (2019), pp. 45–54. DOI: 10.1016/j.mee.2018.12.002.

[241] Rajesh Ghosh et al. "Fabrication of laser printed microfluidic paper-based analytical devices (LP- μ PADs) for point-of-care applications". In: *Scientific Reports* December 2018 (2019), pp. 1–11. DOI: 10.1038/s41598-019-44455-1.

[242] Joshua M Campbell et al. "Microfluidic and Paper-Based Devices for Disease Detection and Diagnostic Research". In: *International journal of molecular sciences* (2018). DOI: 10.3390/ijms19092731.

[243] Mazhar Sher et al. "Paper-based analytical devices for clinical diagnosis: recent advances in the fabrication techniques and sensing mechanisms". In: *Expert Rev Mol Diagn.* 17.4 (2017), pp. 351–366. DOI: 10.1080/14737159.2017.1285228.Paper-based.

[244] Paul Yager et al. "Microfluidic diagnostic technologies for global public health". In: *Nature* 442.July (2006). DOI: 10.1038/nature05064.

[245] Sandeep Kumar Vashist et al. "Emerging Technologies for Testing". In: *Trends in Biotechnology* 33.11 (2015), pp. 692–705. ISSN: 0167-7799. DOI: 10.1016/j.tibtech.2015.09.001. URL: http://dx.doi.org/10.1016/j.tibtech.2015.09.001.

[246] K L M Moran, D Lemass, and R O Kennedy. "Surface Plasmon Resonance e Based Immunoassays : Approaches , Performance , and Applications". In: *Handbook of Immunoassay Technologies*. Elsevier Inc., 2018, pp. 129–156. ISBN: 9780128117620. DOI: 10.1016/B978-0-12-811762-0.00006-2. URL: http://dx.doi.org/10.1016/B978-0-12-811762-0.00006-2.

[247] Joonki Hwang, Sangyeop Lee, and Jaebum Choo. "Application of a SERS-based lateral flow immunoassay strip for the rapid and sensitive detection of staphylococcal enterotoxin B". In: *Nanoscale* 8.22 (2016). DOI: 10.1039/c5nr07243c.

[248] Ying Wang et al. "A SERS-based lateral flow assay for the stroke biomarker S100- β". In: *Microchimica Acta* (2019).

[249] Boris N Khlebtsov et al. "SERS-based lateral flow immunoassay of troponin I by using gap- enhanced Raman tags". In: *Nano Research* 12.2 (2019), pp. 413–420.

[250] Vi Tran et al. "Rapid , Quantitative , and Ultrasensitive Point-of-Care Testing : A Portable SERS Reader for Lateral Flow Assays in Clinical Chemistry". In: *Angewandte Chemie - International Edition Communications* (2019), pp. 442–446. DOI: 10.1002/anie.201810917.

[251] Carly A Holstein et al. "Immobilizing affinity proteins to nitrocellulose : a toolbox for paper-based assay developers". In: *Analytical and Bioanalytical Chemistry* (2015). DOI: 10.1007/s00216-015-9052-0.

[252] Millipore. "Rapid Lateral Flow Test Strips: Considerations for product development". In: *EMD Millipore Corporation, Billerica, MA., Germany* (2008), pp. 1–42. URL: papers3://publication/uuid/F7A438B7-0365-4DA7-83B5-4375A96FD0C4.

[253] George P Smith. "Filamentous Fusion Phage : Novel Expression Vectors That Display Cloned Antigens on the Virion Surface". In: *Science* 508.1984 (1985).

[254] Chien-hsun Wu et al. "Advancement and applications of peptide phage display technology in biomedical science". In: *Journal of Biomedical Science* (2016), pp. 1–14. ISSN: 1423-0127. DOI: 10.1186/s12929-016-0223-x. URL: http://dx.doi.org/10.1186/s12929-016-0223-x.

[255] Rodrigo Barderas. "The 2018 Nobel Prize in Chemistry : phage display of peptides and antibodies". In: *Analytical and Bioanalytical Chemistry* (2019), pp. 2475–2479.

[256] Simon Andrews. *FastQC*. 2019. URL: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

[257] Patrick D Schloss et al. "Introducing mothur : Open-Source , Platform-Independent , Community-Supported Software for Describing and Comparing Microbial Communities". In: *APPLIED AND ENVIRONMENTAL MICROBIOLOGY* 75.23 (2009), pp. 7537–7541. DOI: 10.1128/AEM.01541-09.

[258] Ran Zhao et al. "EBioMedicine Implications of Genetic and Epigenetic Alterations of CDKN2A (p16 INK4a) in Cancer". In: *EBIOM* 8.127 (2016), pp. 30–39. ISSN: 2352-3964. DOI: 10.1016/j.ebiom.2016.04.017. URL: http://dx.doi.org/10.1016/j.ebiom.2016.04.017.

[259] Sanna K. Myöhänen, Stephen B. Baylin, and James G. Herman. "Hypermethylation can selectively silence individual p16(ink4A) alleles in neoplasia". In: *Cancer Research* 58.4 (1998), pp. 591–593. ISSN: 00085472.

[260] Timothy L. Bailey and Charles Elkan. "Fitting a mixture model by expectation maximization to discover motifs in biopolymers". In: *Proceedings of the Second Internation Conference on Intelligent Systems for Molecular Biology* (1994), pp. 3–9.

[261] Hong Lou et al. "Low-cost HPV testing and the prevalence of cervical infection in asymptomatic populations in Guatemala". In: *BMC Cancer* (2018), pp. 1–7.

[262] W Alter M Muller et al. "Light sheet Raman micro-spectroscopy". In: *Optica* 3.4 (2016).

[263] Israel Rocha-mendoza et al. "Rapid spontaneous Raman light sheet microscopy using cw-lasers and tunable filters". In: *Biomedical Optics Express* 6.9 (2015), pp. 3449–3461. DOI: 10.1364/BOE.6.003449.

[264] Qiang Zhang et al. "Design considerations for SERS detection in colloidal solution : reduce spectral intensity fluctuation". In: *Journal of Raman Spectroscopy* December 2015 (2016), pp. 395–401. DOI: 10.1002/jrs.4828.

[265] Renee R Frontiera et al. "Surface-Enhanced Femtosecond Stimulated Raman Spectroscopy". In: *The Journal of Physical Chemistry Letters* (2011), pp. 1199–1203. DOI: 10.1021/jz200498z.

[266] Yashar E Monfared et al. "Continuous-Wave Coherent Raman Spectroscopy via Plasmonic Enhancement". In: *Scientific Reports* August (2019), pp. 1–7. DOI: 10.1038/s41598-019-48573-8.

[267] C L D Lee and K C Hewitt. "First demonstration of surface enhanced- stimulated Raman spectroscopy (SE-SRS) using low-power CW sources". In: *Faraday Discussions* (2017), pp. 227–232. DOI: 10.1039/c7fd00137a.

[268] J.M. Gutierrez Villarreal, J.A. Gaspar Armenta, and L.A. Mayoral Astorga. "Surface plasmon field enhancement : excitation by a short pulse or narrow beam of light". In: *Journal of the Optical Society of America B* 35.5 (2018), pp. 1040–1045.

[269] Martin Lee et al. "In vivo imaging of the tumor and its associated microenvironment using combined CARS / 2- photon microscopy In vivo imaging of the tumor and its associated microenvironment using combined CARS / 2-photon microscopy". In: *IntraVital* 9087 (2015). DOI: 10.1080/21659087.2015.1055430.